

Archiviazione e conservazione delle risorse digitali

Les archives électroniques. Manuel pratique

a cura della Direzione degli Archivi di Francia

(febbraio 2002)

Traduzione

di Paolo Franzese

La progressiva affermazione della descrizione archivistica in formato digitale e la crescente produzione di archivi elettronici accanto a quella di archivi tradizionali pongono gli archivisti davanti all'urgenza di un adattamento di schemi e di metodologie di lavoro alla nuova situazione determinatasi con i profondi cambiamenti tecnologici di questi anni.

Benché non sempre in modo coordinato e spesso senza la necessaria cura per i criteri di normalizzazione, sia gli istituti archivistici che le istituzioni produttrici di archivi hanno trattato o elaborato considerevoli quantità di documentazione e di basi di dati digitali, che rischiano, per la rapida obsolescenza dei programmi con cui sono state prodotte, per la non conformità dei dati a standard condivisi, per il degrado di supporti o semplicemente per l'oblio in cui esse sono via via cadute, di diventare in tutto o in parte inutilizzabili. Ci si accorge in questi anni di aver dato vita a modalità di produzione documentaria di cui non ci si è preoccupati abbastanza di valutare le incognite e di prevenire i rischi. Chi lavora alla tutela della memoria storica, attraverso la digitalizzazione di documenti, attraverso il controllo di archivi in formazione o attraverso l'elaborazione di inventari o anche di banche dati descrittive, accessibili in locale o in linea, vive il disagio di non avere più a disposizione le cognizioni e i mezzi per conservare, per tempi lunghi, quanto è stato prodotto con finalità di ricerca. Il carattere

effimero delle applicazioni sul web riguarda anche le risorse culturali, che invece potrebbero continuare a rappresentare un patrimonio prezioso di sapere e di conoscenza. Si paga oggi in effetti il debito contratto nei confronti di una rivoluzione tecnologica che segue altre finalità ed è alimentata da altri fattori. I ritardi nella formazione di una cultura del digitale e della sua conservazione si riflettono nella scarsa attenzione a mantenere memoria del contesto in cui si è prodotta una determinata documentazione, dei mezzi e dei formati adoperati, delle scelte fatte.

In questi anni si è cominciato a definire un ambiente normativo per la produzione di documenti digitali. Una circolare dell'AIPA del 23 novembre 2000 individua i requisiti legali che il documento informatico deve avere per produrre effetti giuridici e quelli relativi al formato. Una direttiva del CNIPA del 19 febbraio 2004 stabilisce i compiti della figura del responsabile del procedimento di conservazione sostitutiva. Il Codice dell'amministrazione digitale, pubblicato con il decreto legislativo del 7 marzo 2005, ha poi definito le funzionalità che i sistemi di archiviazione e di conservazione dei documenti informatici dovrebbero garantire. A livello internazionale, il Consiglio Internazionale degli Archivi ha presentato il 9 gennaio 2004 all'UNESCO un rapporto sugli archivi digitali e una *Communication on the Digitisation and online accessibility of cultural material and Digital Preservation* è stata pubblicata il 24 agosto 2006 dalla Commission of the European Communities. Da segnalare ancora documenti sulla conservazione digitale prodotti dai [National Archives](#) inglesi nel giugno del 2003 e un [Rapporto sulla conservazione preventiva dei documenti in Finlandia, Francia, Italia, Spagna e Regno Unito](#), recentemente pubblicato da Armida Batori,

Il dibattito e la riflessione su questi temi ha prodotto, in questi anni, risultati considerevoli, anche grazie a progetti internazionali che hanno puntato a studiare le questioni sul tappeto e a progettare standard tecnologici e norme di corretto comportamento. In questo ambito, all'attenzione verso i problemi

informatici e tecnologici si affianca la preoccupazione per la gestione degli archivi digitali e per i rapporti che il loro sviluppo implica con le nuove tecnologie, senza che questi due momenti però riescano adeguatamente a fondersi, in modo da permettere di focalizzare finalità e obiettivi e di elaborare, in tempi non troppo lunghi, strategie efficaci. Questa difficoltà è forse un ulteriore riflesso del non facile rapporto fra discipline diverse, che riescono a trovare punti di contatto più misurandosi nel concreto dei problemi e delle soluzioni da escogitare, che sul piano del confronto teorico. Essa testimonia però anche le modalità con cui sta entrando a far parte del patrimonio professionale degli archivisti italiani e di quello procedurale degli istituti l'applicazione di buone norme riconosciute e da condividere, sia relativamente ai contenuti che alla forma di presentazione delle risorse archivistiche.

Si fa strada tuttavia la consapevolezza che accanto al problema della durata dei supporti e dei costi del riversamento dei dati ci sia quello della conservazione dell'utilizzabilità delle applicazioni informatiche e della leggibilità dei dati, che rimanda alla questione del rispetto di standard descrittivi e gestionali.

Il [*Manuale pratico per gli archivi elettronici*](#) pubblicato nel febbraio del 2002 dalla Direzione degli Archivi di Francia affronta, dal lato degli archivisti, proprio la sfida di una più efficace impostazione dei rapporti fra informatica e archivistica, cercando di individuare e di sottolineare il ruolo degli archivisti all'interno di una nuova cultura della conservazione delle risorse digitali e le relazioni che essi dovrebbero costruire con gli altri specialisti implicati nei sistemi di gestione degli archivi.

La realizzazione del manuale scaturisce in effetti da lavori condotti nell'ambito di un vasto programma governativo in materia di conservazione degli archivi elettronici e da un'intensa collaborazione fra la Direzione degli Archivi di Francia (DAF) ed altri istituti e amministrazioni, i cui risultati si

rilevano dal puntuale ricorso all'esperienza maturata appunto anche da questi interlocutori. Infatti, secondo le stesse intenzioni degli autori, il manuale non si rivolge soltanto agli archivisti, ma anche a tutti coloro a cui spetta assumersi responsabilità nella costruzione di sistemi e di basi di dati.

Il manuale si propone di elaborare concetti e strumenti di analisi degli oggetti e degli obiettivi di strategie efficaci, ma, con un atteggiamento pragmatico, si misura molto frequentemente anche con problemi concreti, il cui esame è sempre accompagnato con l'illustrazione di esperienze. Volendo costituire uno strumento di lavoro, esso illustra non solo istruzioni e procedure, ma presenta anche consigli pratici per far fronte a effettive situazioni di disagio e di difficoltà.

Piuttosto che un punto d'arrivo, l'opera si presenta invece come il punto di partenza di una serie di studi e di pubblicazioni di argomento più specifico, che, annunciati dal manuale, effettivamente cominciano ad essere già disponibili in linea. Fra maggio e agosto del 2006 infatti la Direzione centrale della sicurezza dei sistemi d'informazione (DCSSI) del Segretariato generale della difesa nazionale francese ha pubblicato, attraverso il proprio [sito web](#), diversi studi risultanti dalla collaborazione con la Direzione degli Archivi di Francia, con il Ministero della cultura e della comunicazione e con la Direzione generale per la modernizzazione dello Stato del Ministero dell'economia, delle finanze e dell'industria, sulla base di uno lavoro commissionato ad una società privata. La stessa DAF ha pubblicato nel marzo del 2006 uno [Standard di scambio](#) di dati per l'archiviazione e il successivo 13 settembre ha emanato [istruzioni](#) sulla pubblicazione degli studi commissionati dalla (DCSSI) sull'archiviazione elettronica sicura nel settore pubblico.

Presupposto, ma anche risultato delle argomentazioni presentate dal manuale, è la necessità che i produttori di archivi riconoscano il ruolo dell'archivista, non soltanto alla fine del processo e a conclusione delle procedure di conservazione, ma soprattutto in fase di progettazione e di programmazione.

Spesso percepito come fattore di complicazione, di ritardi e di aggravio di spese, l'intervento del "professionista della conservazione a lungo termine" può rivelarsi invece risolutivo e utile e, allo stesso tempo, capace di produrre conseguenze anche in termini di economicità e di risparmio di risorse. Perché sia veramente così, l'archivista deve però legittimare il proprio ruolo, dotandosi dei necessari strumenti analitici e metodologici, come la capacità di comunicare con le altre figure di esperti, di analizzare i complessi di dati e i sistemi per la loro gestione e descrizione. Anzi uno dei suoi compiti è proprio quello di aiutare il produttore dei documenti a mettere a fuoco le sue intenzioni e le finalità per cui produce i documenti. L'archivista, sottolinea il manuale, non può considerarsi responsabile soltanto dei documenti storici!

All'archivista è assegnata anche la delicata funzione di mediatore fra l'informatico e il pubblico, facendo in modo che, attraverso le opportune soluzioni comunicative, l'utente possa comprendere le strategie stabilite per la gestione dei documenti.

Partendo dalla definizione dell'oggetto stesso dell'archiviazione e dall'analisi delle conseguenze derivanti dalle differenze fra documento cartaceo e documento digitale, il manuale prende in esame soprattutto i problemi di una conservazione sostenibile degli archivi in formazione. Sottolinea quindi la necessità di accompagnare la produzione documentaria con la raccolta puntuale delle informazioni che la riguardano e che vanno a costituire il contesto di creazione, di vita e di trattamento dei documenti. Queste preziose informazioni entrano a far parte dei metadati associati ai documenti, la cui compilazione, secondo standard riconosciuti, non va rinviata al momento del versamento dei dati. A questo proposito, rientra fra i compiti dell'archivista anche la selezione dei dati, con la relativa versione, da conservare in funzione delle esigenze della ricerca storica.

Il manuale sottolinea l'importanza di assicurare una sorta di costante "attenzione tecnologica" ai dati e della garanzia dell'identità e dell'integrità

dei documenti, servendosi delle otto regole definite dal progetto InterPares, ma allo stesso tempo raccomanda pragmatismo e buon senso e competenza dinanzi ai problemi legati alla firma elettronica.

Centrale nel processo di archiviazione è la classificazione del documento, da realizzare secondo un piano generale che non può non scaturire dall'analisi della struttura funzionale del soggetto produttore, senza riflettere necessariamente il suo organigramma. Nell'ambito del manuale, l'archivista è rappresentato come una figura di garante della struttura formale dell'archivio, che ne può assicurare la corretta conservazione a lungo termine.

Il suo ruolo tuttavia acquista davvero un senso soltanto all'interno di uno stretto e funzionale rapporto con il servizio d'archivio dell'istituto produttore, alla cui realtà e ai cui mezzi l'archivista adatta la strategia di conservazione. Anzi è indispensabile che l'istituto riconosca l'utilità di questo obiettivo e dell' "investimento intellettuale, finanziario e umano" che esso implica e che lo inserisca nelle sue linee strategiche.

Anche la tecnica della migrazione, necessaria per la "conservazione a lungo termine", impone decisioni che producono importanti conseguenze, come la scelta fra il cambiamento di supporto fisico e quello riguardante anche il formato o la codifica dei dati, fra l'irreversibilità e la reversibilità del processo.

Il manuale fornisce informazioni e istruzioni sui formati dei documenti, sui criteri di orientamento nelle scelte, sui linguaggi di marcatura (particolarmente raccomandata è la definizione e la memorizzazione di una DTD di XML per le applicazioni su web) e sui supporti, nonché sulle modalità con cui il pubblico può accedere ai dati archiviati.

La strategia di conservazione si riferisce naturalmente anche alla posta elettronica, per la quale deve essere concepito un piano di indicizzazione coerente e funzionale.

Problemi particolari si pongono per l'archiviazione dei siti internet, soprattutto per quanto riguarda le loro applicazioni dinamiche.

Il manuale esamina infine il cruciale problema dei costi dell'archiviazione e della migrazione, di cui prova a fornire l'analisi dei fattori e delle componenti. L'autore della traduzione del manuale si augura di essere riuscito a contribuire, con questo lavoro, allo sviluppo del dibattito in corso e a fornire, in primo luogo agli archivisti, uno strumento utile perché anche in Italia si possa avviare un'efficace strategia d'intervento nel campo della conservazione degli archivi digitali.

Roma, ottobre 2006

Paolo Franzese

Gli archivi elettronici Manuale pratico

Direzione degli Archivi di Francia
(febbraio 2002)

Catherine Dhérent
Conservatore generale

con la collaborazione di

Françoise Banat-Berger (Ministero della Giustizia), **Christian Biard** (SNCF), **Bruno Delmas** (Ecole des Chartes), **Geneviève Drouhet** (gruppo Médéric), **Thibaut Girard** (EDF-GDF), **Claude Huc** (Centro nazionale di studi spaziali), **Marie-Noëlle Leblanc** (Centro degli archivi contemporanei), **Christian Perrot** (SNCF), **Joël Piovre** (Direzione degli Archivi di Francia) e i membri del gruppo di studio della Direzione degli Archivi di Francia sulla conservazione dei documenti elettronici.

Ringraziamo per la rilettura

Claude Aubrie (INRIA, Direzione dell'informazione scientifica e della comunicazione) e **Julien Masanès** (Biblioteca nazionale di Francia)
[non si trascrivono i nomi dei 38 membri del gruppo di studio della Direzione degli Archivi di Francia sull'archiviazione dei documenti elettronici, presieduto da **Martine de Boisdeffre**, direttrice degli Archivi di Francia.]

Sommario

I. Introduzione

II. L'oggetto di archiviazione

III. Ruolo e competenze dell'archivista

III.1. L'archivista e i suoi interlocutori

III.2. I requisiti che deve avere un servizio d'archivio

III.3. Strategie da mettere in opera

IV. Preparare il versamento

IV.1. Reperire i dati digitali esistenti

IV.2. Costituire i metadati

IV.3. Redigere il piano di classificazione del soggetto produttore

IV.4. Redigere i quadri di gestione

IV.5. Selezionare i dati

- Gli obblighi stabiliti dalla CNIL¹

IV.6. Conservare i dati in formato cartaceo, digitale o analogico

- Bisogna digitalizzare i documenti cartacei?

IV.7. Garantire l'identità e l'integrità dei dati elettronici

Riflessioni sulla firma elettronica

IV.8. Rendere sicuro il trasferimento

V. Garantire la conservazione

V.1. Condizioni minime richieste per l'accettazione di un versamento di dati elettronici

V.2. Scelta dei formati e dei supporti

V.2.1. Formati e linguaggi di codificazione

- I formati testuali
 - I trattamenti del testo
 - I metalinguaggi e i linguaggi di marcatura
- I formati immagine
- I formati misti (testo e immagine)
- Criterio di scelta dei formati

V.2.2. Supporti

V.3. La conservazione a lungo termine dei dati

V.3.1 Migrazione dei dati

V.3.2. Le microforme COM (Computer Output Microform)

V.4. I locali di archiviazione

VI. Mettere a disposizione i dati

VI.1. Gli strumenti di ricerca

¹ Commission Nationale Informatique et Libertés (<http://www.cnil.fr>).

VI.2. Comunicazione dei dati**VII. Studi di casi****VII.1. Basi di dati****VII.2. I documenti provenienti dalla burotica² e la posta elettronica****VII.2.1. I documenti provenienti dalla burotica****VII.2.2. La posta elettronica****VII.3. Risorse del Web****VIII. Conclusioni****IX. Allegati****IX.1. Peso medio dei documenti elettronici e capacità dei supporti****IX.2. I mestieri legati all'informatica: esempio del Ministero dei lavori pubblici****IX.3. Alcuni principi di base per valutare il costo dell'archiviazione elettronica****IX.3.1. La cattura****IX.3.2. L'archiviazione****IX.3.3. La consultazione/archiviazione****IX.3.4. La migrazione****X. Glossario delle sigle****XI. Bibliografia****XI.1. Norme****XI.2. Idee generali****XI.3. Metadati****XI.4. Archiviazione di basi di dati****XI.5. Archiviazione delle risorse Web**

² Complesso delle tecnologie, specialmente informatiche, che permettono l'automatizzazione del lavoro d'ufficio.

1. Introduzione

L'8 e il 9 marzo 2000, la Direzione degli Archivi di Francia ha organizzato delle Giornate internazionali sulla conservazione a lungo termine dei documenti elettronici. Questi incontri, nel corso dei quali diciannove esperti del mondo intero hanno esposto le loro esperienze, i loro successi e i loro problemi, hanno avuto un grande pubblico. Vi hanno assistito cinquecento persone provenienti tanto dal settore pubblico che dal settore privato. Nel corso di questi scambi, è emerso il bisogno per i professionisti degli archivi e per i loro interlocutori di disporre di uno strumento di riferimento semplice, di primo approccio, che offra loro raccomandazioni e procedure generali e redatto in una lingua comprensibile per loro.

Il contesto è oggi più favorevole rispetto a dieci anni fa. Lavori francesi, europei e internazionali nel campo dell'interoperabilità, degli standard di scambio e della conservazione a lungo termine confortano le azioni degli archivisti e dei gestori di documenti. Si possono citare tra essi gli standard internazionali validati dal W3C (World Wide Web Consortium) per la strutturazione e lo scambio dei documenti, il manuale e l'appello all'industria lanciato dal DLM-Forum³ a livello europeo, infine il programma d'azione governativo francese (PAGSI⁴), che ha avuto come conseguenza in particolare la creazione nel 1998 della MTIC⁵, divenuta nel 2001 ATICA⁶. Si può anche citare la circolare del Primo ministro del 2 novembre 2001 relativa alla gestione degli archivi nei servizi e negli stabilimenti pubblici dello Stato, che sottolinea il ruolo dei responsabili di archivi intermedi nella buona gestione e nella comunicazione dei dati elettronici.

La Direzione degli Archivi di Francia è stata associata a tutti i lavori condotti nel quadro di questo programma governativo per quanto riguarda la conservazione a lungo termine dei documenti elettronici: la preparazione della Guida pubblicata dalla MTIC nel gennaio 2001, le riflessioni del gruppo incaricato del lavoro preparatorio e di un rapporto sullo stabilimento e sulla conservazione degli atti autentici elettronici, in vista della redazione del

³ Istituito nel 1994 su richiesta del Consiglio dell'Unione Europea, è un forum multidisciplinare che si occupa dei problemi del management, della memorizzazione, della conservazione e del recupero di dati *machine/readable*. Al suo interno collaborano pubbliche amministrazioni, servizi archivistici nazionali, rappresentanze di imprese e di istituti di ricerca.

⁴ Programme d'action gouvernemental pour la société de l'information.

⁵ Mission pour l'introduction des technologies de l'information et de la communication dans l'administration.

⁶ Agence pour les technologies de l'information et de la communication dans l'administration.

decreto previsto dalla legge del 13 marzo 2000⁷, le riunioni del Consiglio superiore del notariato.

Non è più il tempo in cui, giuridicamente, la carta era il solo vettore affidabile per trasmettere alla posterità dati e informazioni: il diritto internazionale e il diritto europeo conferiscono ormai, a certe condizioni, lo stesso valore allo scritto elettronico e al documento su carta.

Se si poteva ancora, fino a poco tempo fa, ripararsi dietro il “diritto alla dimenticanza” per lasciar sparire dei dati informatizzati, la legge concilia ora, nel modo più netto, protezione della vita privata e bisogni della ricerca storica, il che mette il contenuto delle memorie dei computer sullo stesso piano giuridico dei fondi archivistici tradizionali. Insomma, la conservazione a lungo termine e la restituzione di tutto ciò che è prodotto in modo digitale e registrato su supporto magnetico o ottico sono ormai temi di cui nessun archivista ha il diritto di disinteressarsi.

Frutto di perizia e di esperienze diverse, questo manuale è stato concepito per archivisti, ma il suo contenuto potrà essere utile a tutte le persone capaci di intervenire nella costituzione dei sistemi e delle basi. La sua ambizione è in effetti d'aiutare gli uni e gli altri ad assicurare la preservazione e l'accessibilità a lungo termine dei documenti e dei dati elettronici, soprattutto indicando loro ciò che è attualmente possibile mettere in pratica, tenuto conto dello stato delle tecniche e della regolamentazione.

Questo manuale, destinato a trovarsi nell'ufficio di ogni gestore di archivi contemporanei, non è tuttavia una raccolta di raccomandazioni tecniche e funzionali, per le quali invece noi consigliamo ai lettori di consultare documenti come la norma AF-Z 42-013⁸, versione 2 del dicembre 2001, e la *Guida* pubblicata dalla MTIC.

La preoccupazione di essere pragmatici ha spinto a operare, quando necessario, una distinzione fra molti casi o molti scenari: si troveranno così nel presente lavoro sia le procedure ideali che consigli per far fronte a situazioni di degrado, per ritrovare degli arretrati o per salvare ciò che può esserlo di sistemi mal concepiti sin dall'origine.

Questo manuale sarà completato da pubblicazioni più specifiche, ciascuna adatta a un modo di produzione o di restituzione dell'informazione digitale.

⁷ Legge sull'adattamento del diritto della prova alle tecnologie dell'informazione e sulla firma elettronica.

⁸ “Archiviazione elettronica. Specifiche relative alla concezione e allo sfruttamento di sistemi informatici per assicurare la conservazione e l'integrità dei documenti immagazzinati in questi sistemi (organizzazione e gestione dei documenti d'archivio. *Record management*. Gestione e conservazione dei documenti in sicurezza)”.

La presente versione è destinata a essere sostituita da una versione regolarmente aggiornata che si troverà in formato elettronico sul sito della Direzione degli Archivi di Francia (<http://www.archivesdefrance.culture.gouv.fr>).

II. L'oggetto di archiviazione

In ambiente digitale, l'unità elementare d'informazione è il dato, che è soltanto una catena di caratteri o byte costituiti da bit (0 o 1). I dati sono scritti su supporti digitali. È indispensabile ricorrere a strumenti tecnologici per gestirli e per leggerli. In questo, essi differiscono interamente dal documento cartaceo che veicola un messaggio immediatamente utilizzabile.

In ambiente cartaceo, i dati e la loro formattazione e la loro messa in opera si trovano sullo stesso supporto e sullo stesso documento. L'archivista dispone, sullo stesso documento, di una gran parte degli elementi che serviranno alla sua critica, al suo studio diplomatico. Del tutto diversa è la situazione in ambiente elettronico: le due cose sono separate, conservate spesso a cura di servizi differenti. Esse devono tuttavia essere riunite insieme per rendere i dati comprensibili e idonei a costituire l'oggetto di un'analisi storica.

Una tipologia sommaria dei differenti oggetti di archiviazione dimostra le loro grandi differenze e specificità. L'archivista dovrà conservare nel lungo termine tre grandi tipi di oggetti digitali:

- il documento che rende conto di un'attività o di un ragionamento, i cui limiti sono fissati e precisati e la cui misura è in genere affidabile (esempio: documento gestito con trattamento di testo, posta elettronica);
- la base di dati che è una riserva di dati a volte molto voluminosa, dinamica, periodicamente aggiornata;
- alcuni oggetti d'informazione digitali che associano i due tipi precedenti sono composti da elementi statici e da elementi dinamici e contengono legami con risorse esterne (esempio: siti web).

Si può poi analizzare questi grandi tipi e tener conto di altri elementi d'identità, sapendo per esempio:

- che una base di dati può comportare molti documenti;
- che il documento può essere prodotto in un processo di *workflow* (flusso di lavoro che regola lo svolgimento di un'azione) e avere diverse versioni prima della validazione;
- che un nuovo oggetto d'informazione digitale può risultare dalla riorganizzazione, dalla selezione e perfino dalla ricerca di dati.

I dati e i documenti digitali sono certamente più facili da manipolare, duplicare, trasferire. Ma essi sono ugualmente più facili da modificare e possono molto rapidamente e in modo irreversibile perdere la loro integrità d'origine.

Esigenze nuove, dovute alla rapida evoluzione dei linguaggi, dei formati e dei supporti digitali, si presentano dunque per la conservazione a lungo termine dell'informazione.

I dati e i documenti digitali sono accessibili soltanto dopo che sia terminato il lavoro di descrizione, di valutazione e di selezione. La descrizione è tanto più difficile quanto più i contenuti sono compositi e quanto più coesistono con tipi documentari su carta. La loro classificazione è tanto più cruciale per la ricerca dell'informazione quanto più la loro produzione è abbondante. Essi sono differenti dai documenti su carta, la cui consultazione non è condizionata dall'ultimazione del lavoro di classificazione e di descrizione.

Inoltre un'applicazione non sostituisce necessariamente un tipo documentario esistente in formato cartaceo. Essa può anche raggruppare molti tipi documentari preesistenti su carta, essere composta da elementi di uno o più tipi documentari su carta, ma anche da dati nuovi che non costituivano anteriormente oggetto di trascrizione. Il tipo documentario elettronico infine può essere interamente nuovo nel suo contenuto informativo, perché risulta dalla constatazione di lacune nei tipi documentari su carta preesistenti.

Così la Nouvelle Chaîne Pénale di Parigi e della regione parigina, nella quale sono registrati gli affari portati davanti alle sezioni penali dei tribunali di prima istanza fin dal 1975, integra informazioni che prima davano luogo alla creazione di molti tipi di documenti differenti e conta una dozzina di moduli funzionali gestiti con 265 tavole, di cui 5 sono le principali.

È dunque con una grande varietà di oggetti digitali, i cui contenuti si sono evoluti in relazione ai documenti cartacei preesistenti, che l'archivista deve confrontarsi. A ciascuno di questi tipi corrisponde una specifica metodologia di conservazione.

III. Ruolo e competenze dell'archivista

L'archivista abituato a oggetti stabili e perenni si trova di fronte ormai a un compito più complesso, del quale non percepisce ancora tutta l'estensione.

Descrivere dati digitali, valutarli e selezionarli prende molto più tempo di quello che richiede classificare archivi in formato tradizionale o redigere uno strumento di ricerca di tipo classico, anche se molto dettagliato.

Da questo deriva la necessità di avere un metodo globale di reperimento dell'informazione, dei suoi circuiti e della produzione documentaria. L'archivista ha il dovere di intervenire molto più rapidamente nel ciclo di vita dell'informazione. Egli continua a essere in modo privilegiato e quasi esclusivo responsabile dei documenti quando questi hanno raggiunto l'età storica o definitiva. Ma deve anche rivendicare un ruolo a monte della loro catena di trattamento.

Nell'impresa come nell'amministrazione, in effetti, gli sconvolgimenti frequenti di organigrammi e il rapido rinnovamento dei responsabili privano sempre più la tradizione orale della sua efficacia e rendono sempre più necessaria un'organizzazione dei flussi documentari a partire dall'origine: produrre documenti dicendo che per la conservazione si vedrà più tardi è una politica di corta visuale, incompatibile con la ricerca crescente di qualità e con la necessità frequente di giustificare le decisioni precedenti. Questo è tanto più vero a proposito del documento elettronico. L'unica soluzione è far intervenire a monte, al fianco dei produttori e dei tecnici, l'archivista che sa ciò che bisogna conservare e come.

Questo nuovo ruolo dell'archivista è perfettamente affermato nella norma ISO 15489 sul *Record Management*. Tuttavia secondo la misura e il tipo di istituzione da cui dipende, l'archivista s'impegnerà a livelli diversi all'inizio del ciclo di vita documentaria. Questo è possibile se gestisce gli archivi di un comune o di un'impresa e ha il compito molto preciso per questa produzione di essere associato strettamente all'elaborazione delle strutture di file, delle strutture ad albero, all'assegnazione del nome dei repertori, alla presa in conto della valutazione, della revisione della valutazione, dei diritti e della comunicabilità, prima che i dati stessi siano creati, mentre invece sarà impossibile per il responsabile di un servizio di archivi dipartimentali, responsabile della gestione a lungo termine delle informazioni prodotte da centinaia di soggetti differenti. Quest'ultimo si applicherà in tal caso a formare buone testimonianze e ottimi corrispondenti in ogni istituto e a farli integrare nella catena di *Record management* di questa.

L'archivista è in grado di rispondere ai nuovi bisogni perché le regole e i metodi che mette in opera per trasmettere alle future generazioni i documenti di formato tradizionale non sono per niente rimessi in discussione in un contesto di produzione documentario dominato dall'informatica. Egli dovrà tuttavia richiamare e far riconoscere queste competenze anche in materia di documenti elettronici, perché i diversi soggetti che intervengono nella

creazione e nell'analisi dei dati e dell'architettura dei sistemi d'informazione, non pensano sempre a far appello alle capacità degli archivisti, di cui a volte non suppongono neppure l'esistenza. È vero che l'intervento di un professionista della conservazione a lungo termine può essere percepito come una complicazione supplementare, suscettibile di ritardare la realizzazione dei progetti e di generare un sovrappiù di spese. In un tale contesto, l'archivista avrà dunque da dimostrare la legittimità della sua presenza su questo terreno, soprattutto spiegando che formalizzare i sistemi di gestione e di produzione documentaria si rivela molto vantaggioso a lungo termine e ricordando che perdere documenti a causa di un'archiviazione difettosa o trovarsi nell'incapacità di produrli in caso di contenzioso può, nel senso proprio del termine, rivelarsi rovinoso.

Questa necessità di perorare la propria causa non è il solo ostacolo che l'archivista rischia di incontrare sulla sua strada, perché bisogna tener conto di un'altra difficoltà: l'utilizzazione delle stesse parole con significati differenti, che complica a volte il dialogo fra persone di mestieri differenti e rende certi ambienti professionali (gli informatici, per esempio) reticenti alla collaborazione con gli archivisti o scettici quanto all'opportunità di questa. Così avviene con la parola *archiviazione* che l'archivista interpreta come il trasferimento di documenti che hanno cessato di essere di utilità corrente verso un locale di conservazione e/o un servizio d'archivio che ha la capacità di riceverli, quando per altri significa "pulizia" dei sistemi e delle memorie di tutto ciò che non è più d'uso corrente, tramite lo stoccaggio su altri supporti d'accesso problematico, per non dire ipotetico (il che spinge a volte *archiviazione* a significare praticamente *distruzione!*). Altrettanto con la parola *archivio* – che gli archivisti privano raramente della *s* finale⁹ – significa servizio d'archivio presso i bibliotecari e i documentalisti e sistema o servizio di archiviazione virtuale presso alcuni informatici. *Documentazione*, infine che per gli archivisti designa sia una disciplina e una professione in rapporto con la ricerca e la restituzione dell'informazione, sia un insieme di informazioni riunite su un dato oggetto, è considerata da molti informatici come l'equivalente di *metadati*. L'archivista farà in modo dunque di essere perfettamente compreso da questi interlocutori e si impegnerà a far esplicitare da questi le loro proprie terminologie.

L'archivista deve badare a quali sono i suoi ruoli e attribuzioni esatti in materia di documenti elettronici. A lui spetta:

- identificare le esigenze in materia di conservazione;
- aiutare i produttori nella creazione e nella sistemazione dei dati e dei documenti in modo che essi abbiano le seguenti caratteristiche: autenticità, affidabilità, integrità, capacità di essere riutilizzati;
- aiutare nella modellizzazione dei circuiti dell'informazione dell'istituto;

⁹ In francese la *s* forma il plurale.

- spingere i produttori a conservare documenti di lavoro e stati preparatori non convalidati, ufficiosi, ma che permettono di comprendere la genesi dei fascicoli e costituiscono tracce storiche importanti;
- aiutare nell'installazione di un sistema di conservazione affidabile, integro, conforme alle esigenze del contesto regolamentare;
- fare in modo che tutti i documenti siano integrati nel sistema di conservazione installato;
- assicurarsi che il sistema di conservazione sia correttamente mantenuto e che sia perenne;
- assicurare gli accessi e le possibilità di ritrovare i dati in funzione di diritti determinati (soprattutto dalla legislazione in vigore).

L'archivista dispone perciò di qualità risultanti dalla sua formazione:

- capacità di comunicare con altri specialisti;
- capacità di analizzare sistemi complessi e di descriverli;
- capacità di valutare i documenti e di descriverli.

L'archivista è un attore, ma è neutrale e il solo della catena di trattamento documentario preoccupato del tempo e dello spazio, perché da un lato considera l'utilizzo del documento da parte di altre comunità diverse da quella dei produttori, dall'altro si preoccupa della conservazione dei dati oltre la loro temporanea utilità immediata.

Se al momento della creazione dei dati egli deve a sé stesso di essere il più attivo e intraprendente, in seguito egli dovrà resistere alla tentazione di ricostruire l'informazione. In effetti, laddove la produzione tradizionale dell'informazione in formato cartaceo era relativamente inquadrata e strutturata, quella dell'informazione elettronica si è sviluppata in modo selvaggio al di fuori dei circuiti esistenti o ufficiali. È così che l'archivista può trovarsi confrontato con produzioni burotiche per le quali nessuna salvaguardia è stata prevista, dove si trovano X versioni di un documento senza poter determinare quella definitiva. L'archivista allora potrebbe essere tentato di compensare al momento dell'archiviazione definitiva e di fare ciò che non è stato fatto durante il ciclo di vita dell'informazione. Vista l'immensità del compito al quale allora si troverebbe davanti, è sconsigliabile provare a ricostituire da sé un'informazione mal documentata e dunque a volte impossibile da conservare. Meglio sarà allora non archivarla affatto.

Per certi dati particolarmente importanti, ciò sembrerà forse indispensabile o sarà chiesto dal soggetto produttore o detentore dell'informazione. Ma questa ricostruzione della documentazione e dei dati dovrà essere eccezionale, dovrà sempre essere fatta insieme con il loro produttore iniziale e dovrà essere

convalidata da questo allo scopo di evitare ogni soggettività da parte dell'archivista, che non conosce totalmente il contesto di produzione d'origine.

Bisogna dunque domandarsi qual è la pertinenza dell'informazione incontrata al momento dell'archiviazione. Se il quadro nel quale essa è stata prodotta non permette di conservarla nella sua totalità in modo soddisfacente, bisogna esaminare se essa può esserlo in modo parziale e se si dispone di elementi sufficienti per spiegare il suo stato lacunoso al futuro ricercatore nella descrizione; il lettore potrà così valutare quale peso accordare all'informazione.

III.1. L'archivista e i suoi interlocutori

Più che in passato, l'archivista dovrà dunque partecipare all'elaborazione e alla diffusione delle politiche, pratiche e linee direttive dell'istituto relative alla conservazione dei documenti nel rispetto della legislazione applicabile agli archivi pubblici e privati.

Ma da solo egli non può fare gran cosa. Salvare i documenti e le informazioni prodotte in modo digitale da una distruzione sconsiderata e assicurare la loro accessibilità a lungo termine non può essere che un'opera collettiva. Ciascuna delle parti interessate alla creazione e alla vita delle basi, dei siti e dei sistemi deve apportare il suo contributo al processo di salvataggio e di archiviazione.

Titoli e attribuzioni di queste parti variano secondo i soggetti e le circostanze. L'archivista dovrà infatti identificare alcune grandi funzioni e i loro titolari (amministrazione dei dati, creazione dei diritti, gestione degli spazi di lavoro, creazione dei dati, salvaguardia, gestione delle reti, sicurezza ...). È con essi che troverà le diverse applicazioni digitali e assicurerà il loro trattamento. Questi interlocutori sono:

- quelli che producono dati e informazioni (i produttori e/o i “*maitres d'ouvrage*”);
- quelli che li amministrano (amministratori di dati, professionisti della qualità¹⁰), li trattano e li diffondono (documentaristi, corrispondenti dei servizi di archivi storici)
- Quelli il cui mestiere è di concepire e di mettere in opera, per i precedenti, le soluzioni e le infrastrutture necessarie, gli informatici.

¹⁰ Finalità dei “*qualiticiens*” è migliorare la qualità alleggerendo i vincoli all'interno delle organizzazioni. Questi professionisti si propongono di mettere la soddisfazione del cliente al centro della strategia dell'impresa, di condividere e perpetuare il saper fare all'interno di un'organizzazione, di sostenere il trasferimento delle conoscenze, di identificare i fattori di anomalie, di assicurare la perfetta tracciabilità a tutti i livelli.

L'archivista sarà così attento al fatto che per una stessa applicazione egli non può aver a che fare con più produttori, con più informatici, interni o esterni all'istituto produttore dei dati.

A volte queste diverse funzioni sono identificate, durante la conduzione di un progetto, con le espressioni “maîtrise d'ouvrage” e “maîtrise d'oeuvre”¹¹. Questi concetti in uso in materia di costruzioni possono essere trasposti nella conduzione dei progetti informatici, perché riguardano attori di specialità differenti, la cui qualità del lavoro di collaborazione è determinante per l'esito e per la riuscita di operazioni complesse e onerose.

La “maîtrise d'ouvrage” raggruppa i soggetti all'origine della commessa per opposizione alla funzione di fornitore assolta dalla “maîtrise d'oeuvre”. Il “maître d'ouvrage” ha il potere di decisione e di impulso nella conduzione del progetto; in certi istituti, può essere il direttore in persona.

Nei progetti importanti, la “maîtrise d'ouvrage” è organizzata intorno a un direttore o a un capo di progetto che è un esperto del settore funzionale coperto e forgia lo scheletro dell'applicazione, in stretto legame con il “maître d'oeuvre” e con gli utilizzatori. Egli gestisce l'organizzazione del progetto e le linee di bilancio che gli sono proprie.

Il “maître d'oeuvre” del progetto concepisce tecnicamente l'applicazione (traduce le modalità funzionali in linguaggio di programmazione, mentre fa veramente valere le possibilità e gli obblighi tecnici al “maître d'oeuvre” affinché egli adatti al bisogno il progetto funzionale), la sviluppa. Un capo di progetto della “maîtrise d'oeuvre” è designato al livello di ciascun progetto per essere l'interlocutore unico del capo del progetto della “maîtrise d'ouvrage”.

La definizione preliminare del ruolo di ciascuno permette di evitare le interferenze, le ambiguità e consente un migliore scambio nel rispetto delle specialità di ciascuno.

Molteplicità dei partner: esempio di un'applicazione informatica del Ministero della giustizia

Quest'applicazione (progetto) mira a registrare e ad assicurare la continuità degli affari passati davanti ai giudici di applicazione delle pene (che appartengono alla giurisdizione e si situano nel campo dell'applicazione delle pene) e ai loro partner dei SIP (servizi d'inserimento e di probazione, creati fin dal 1999 che si situano nell'amministrazione penitenziaria e sostituiscono gli antichi comitati di probazione e di assistenza ai liberati –

¹¹ Le due espressioni, di cui più avanti si spiega il significato, indicano rispettivamente l'equipe degli esperti del lavoro nella sua fase progettuale e quella degli esperti nella fase finale dell'opera.

ambiente aperto – e i servizi socio-educativi degli istituti penitenziari). Per realizzare un tale progetto, molti interlocutori sono presi in considerazione:

- *la “maîtrise d’ouvrage” dell’amministrazione penitenziaria (lato SIP) e l’ufficio in carica dell’inserimento della popolazione penale;*
- *la “maîtrise d’ouvrage” dei servizi giudiziari (lato JAP) e l’ufficio incaricato dell’informatizzazione delle giurisdizioni (che esso stesso lavora in associazione con altre direzioni o uffici del ministero “competenti” o “esperti”, soprattutto in questo esempio di altri servizi della direzione dei servizi giudiziari come la direzione degli affari criminali e della grazia);*
- *la “maîtrise d’oeuvre” per l’amministrazione penitenziaria e l’ufficio incaricato dell’informatizzazione dei servizi dell’amministrazione penitenziaria all’interno della direzione dell’amministrazione penitenziaria;*
- *la “maîtrise d’oeuvre” per i servizi giudiziari e uno degli uffici della sottodirezione dell’informatica questa volta all’interno della direzione dell’amministrazione generale e delle attrezzature.*

A questo bisogna aggiungere, lato “maîtrise d’oeuvre”, all’interno della sottodirezione dell’informatica, l’ufficio incaricato delle reti e dell’architettura tecnica.

All’interno della “maîtrise d’oeuvre”, appaiono inoltre delle persone-risorse incaricate della redazione del piano di assicurazione qualità, a cui conviene avvicinarsi nella misura in cui queste persone hanno soprattutto incarico della gestione della documentazione della futura applicazione e alle quali si può apportare un aiuto in materia di strutturazione di questa informazione, di cui si dovrà assicurare a termine l’archiviazione.

L’archivista può chiedere di essere associato a alcune di queste funzioni e di seguirle. Può allora partecipare alla redazione di un quaderno dei carichi per la realizzazione di applicazioni e sorvegliare l’installazione dei moduli di archiviazione, partecipare al ricevimento delle bozze, assicurare la formazione dei servizi che utilizzano le applicazioni per la parte archiviazione, badare, insieme con gli informatici durante i trasferimenti nei servizi d’archivio, che i dati migrati non abbiano subito alterazioni e abbiano conservato il loro carattere autentico.

L’archivista può anche apportare il suo aiuto ai servizi per l’installazione di piani di classificazione e di strutturazione di repertori elettronici.

III.2. Le caratteristiche che deve avere un servizio d’archivio

Se l’archivista deve tenersi al corrente delle evoluzioni tecnologiche relative al settore dell’archiviazione dei dati digitali, non può controllarli completamente. Sarebbe particolarmente utile che egli potesse costituire presso di sé un gruppo dedicato alla conservazione dei documenti elettronici. L’ideale sarebbe lo sviluppo di un polo di competenza di questo tipo con un tecnico informatico, come avviene nei grandi servizi di conservazione

nazionali, per esempio il gruppo CONSTANCE al Centro degli archivi contemporanei (CAC) di Fontainebleau.

Quando non è possibile costituire un gruppo di specialisti, l'archivista dovrà trovare un agente del suo servizio, le cui competenze in questo campo possano essere sviluppate, grazie a una formazione specifica.

Altrimenti si ricorrerà ai servizi informatici più vicini, per esempio, i servizi informatici del dipartimento per archivi dipartimentali.

L'archivista dovrà anche aggiungere le competenze di un giurista, sia all'interno, sia facendo ricorso a chi sarà più vicino al suo istituto (oggi la maggior parte delle grandi amministrazioni dispone di un tale servizio).

III.3. Strategie da mettere in opera

Non può esserci buona politica di archiviazione senza una presa di coscienza della sua importanza e una convalida al livello più alto della gerarchia dell'istituto produttore, al livello della direzione generale. Il primo requisito a cui rispondere è dunque il riconoscimento, a questo livello, del carattere strategico di una politica mirante a conservare a lungo termine i dati digitali, per esempio considerando la loro archiviazione nello schema direttivo dell'istituto.

Per convincere i responsabili dell'istituto di produzione e la sua gerarchia dell'utilità della conservazione a lungo termine dei dati digitali e dell'investimento intellettuale, finanziario e umano che essa presuppone, bisogna proporre argomenti di conservazione (giuridica, finanziaria), ma anche di valorizzazione e di riutilizzazione con la comunicazione.

Un'esperienza del Centro degli archivi contemporanei

Il Centro degli archivi contemporanei (Fontainebleau) ha potuto ricavare delle conclusioni positive da un'operazione di salvataggio di inchieste antiche, ma interessanti durante la quale è stato necessario ricostruire tutta la documentazione tecnica che descrive la struttura dei file.

Il compito fondamentale del servizio versante era la pubblicazione di risultati a partire da queste inchieste. I file erano uno strumento di lavoro che aveva valore solo per le pubblicazioni alle quali permetteva di accedere. In questa logica, il produttore dei dati non vi ha prestato attenzione finché esse erano utilizzabili e, una volta utilizzate, non le ha considerate come delle informazioni che hanno valore in sé e la cui conservazione era necessaria. Il CAC ha reso conto della sua operazione di archiviazione, del suo laborioso lavoro di ricostituzione della documentazione associata e ha formulato l'idea che la

conservazione di questa poteva presentare un interesse. Il servizio versante ha così preso coscienza che la perdita dei dati poteva essergli pregiudizievole quanto le inchieste potevano ripetersi nel tempo e che può essere utile accedere a una precedente versione elettronica. Si è dunque deciso di prevedere la costituzione di un comitato di archiviazione che studiasse, tra le altre cose, le misure necessarie per organizzare la conservazione delle inchieste su una lunga durata.

Il ricorso ai dati per esigenze giuridiche può anche essere un argomento in favore di una buona conservazione dei dati elettronici (è così che uno dei responsabili dell'informatizzazione delle giurisdizioni al Ministero della giustizia si è reso conto dell'importanza delle procedure di conservazione perenne quando c'è stato bisogno di ritrovare dei dati per l'affare delle "scomparse dell'Yonne"¹²; per la stessa inchiesta giudiziaria, la SNCF¹³ ha dovuto ritrovare rapidamente gli orari dei treni dell'epoca per verificare la validità delle informazioni fornite dai sospetti ...).

Esempio di strategia del Ministero della giustizia

L'implicazione dell'archivista è cominciata con l'accettazione da parte dei suoi interlocutori della sua presenza all'interno dei comitati di controllo civile e penale del ministero, con l'appoggio del direttore. Questa partecipazione ha permesso l'avvicinamento ai servizi statistici del ministero, i cui obiettivi sono di approntare dei moduli statistici efficaci all'interno degli applicativi, da cui la necessità di intervenire a monte, esattamente come gli archivisti. È d'altronde la lettura del rendiconto del comitato di programmazione della statistica che ha permesso di sapere che il repertorio generale civile nelle giurisdizioni era in corso di rifacimento (per far in modo che tutti gli affari che passano da una giurisdizione, provenienti dal giudiziario o dal ruolo di deposito e di registrazione della cancelleria, siano registrati nel repertorio generale). Uno studio allora ha potuto essere redatto concernente l'archiviazione elettronica dei dati che figurano nel repertorio generale (uno studio identico era stato preparato per le statistiche) che elenca i dati essenziali e precisa che essi sarebbero trasferiti regolarmente all'amministrazione degli archivi. Da questo esempio si capisce che si tratta di intervenire molto a monte: in questo caso specifico, le raccomandazioni del repertorio saranno trasmesse alle società che gestiscono le applicazioni civili installate nei corsi, nei tribunali e nei consigli dei probiviri per essere in seguito considerate e permettere l'evoluzione dei prodotti. In ogni stato di causa quest'evoluzione non si verificherà prima del 2002, 2003 e perfino 2004 (per i tribunali d'istanza) tanto sono grandi i pesi riguardanti gli affari, la redazione dei quaderni dei carichi e le operazioni a valle (test, siti pilota, installazione progressiva ...). Si comprende così che intervenire una volta a giochi fatti (cioè quando gli applicativi sono stati modificati) è controproducente.

La partecipazione a questo comitato ha permesso di posizionarsi durante il lancio di un nuovo progetto di apertura di una "nouvelle chaîne penale" per tutti i tribunali – fuori Parigi e la regione parigina. Dalla fase di redazione del quaderno degli incarichi per

¹² Notizia di cronaca nera in primo piano nel dicembre del 2000 in seguito all'arresto di un uomo, presunto responsabile dell'uccisione di sette ragazze, portatrici di handicap, scomparse fra il 1977 e il 1979.

¹³ Société Nationale des Chemins de Fer.

l'assistenza alla "maîtrise d'ouvrage", l'archivista è stato sollecitato in quanto maestro d'opera associato, la collaborazione con il gruppo incaricato del progetto su tre punti: intervento nei gruppi di lavoro sul recupero dei dati delle antiche applicazioni, valendosi del fatto che sarebbe utile contrassegnare gli affari sufficientemente antichi per essere archiviati, allo scopo di immagazzinarli in una base archivi, evitando così un recupero dei dati nella nuova applicazione troppo importante; definizione, modulo funzionale per modulo funzionale, dei dati destinati a essere perpetuati e dunque a essere riversati in una base archivi centralizzata (ai lati dell'Infocentro, essendo messo in opera lo stesso meccanismo nei due casi di estrazione e di immissione dei dati); consigli sui formati, le trappole da evitare per la digitalizzazione di certi documenti e più generalmente la gestione elettronica di documenti.

In seguito, questa prima integrazione ha permesso all'archivista di essere associato ai lavori di preparazione del nuovo schema direttivo informatico. Quest'associazione è consistita nell'intervenire nel sistema di riferimento tecnologico, facendovi integrare gli strumenti d'interfaccia per l'archiviazione; nell'elaborazione del sistema di riferimento dei mestieri dell'informatica nel ministero facendovi integrare l'archivista in quanto, da una parte, "maître d'ouvrage" associato e, d'altra parte, aiuto per i servizi nei loro metodi di lavoro (allo scopo di migliorare la gestione documentaria); per la preparazione di una metodologia tipo per il lancio di un progetto che segue il livello del progetto, tentando di definire il ruolo di un "maître d'ouvrage" associato; per la cartografia delle applicazioni integrando una colonna "modulo di archiviazione" che permette di distinguere le applicazioni per le quali conviene prevedere un'archiviazione, da quelle per le quali era inutile (controllo dei planning del personale, per esempio); per la revisione del dizionario dei dati suddivisi che permette, per tutte le applicazioni, di intendersi sugli stessi concetti e infine per la definizione dei domini d'azione del ministero (in legame con i mestieri e di conseguenza con le applicazioni), campo sul quale il valore aggiunto dell'archivista era molto reale in ragione della sua conoscenza globale delle attribuzioni del ministero.

Altro esempio del Ministero della giustizia

L'associazione ai lavori di un gruppo di lavoro istituito nel quadro della R.T.T., sui modi di lavorare meglio con i nuovi strumenti (lavoro cooperativo, registrazione della posta, archiviazione della posta elettronica, messa in rete all'interno dell'ufficio ...). Il punto di vista dell'archivista ha potuto così essere inteso e si è intesa soprattutto l'importanza di tener conto dell'archiviazione molto veloce nella catena del trattamento documentario. A partire dalle conclusioni di questo gruppo di lavoro, sono stati creati due altri gruppi. Il primo riguarda l'elaborazione di uno strumento di registrazione della posta unico per l'insieme del ministero (che permetta di assicurare la tracciabilità del controllo della posta e dei documenti del ministero): in quest'ottica, è possibile orientarsi verso un'organizzazione ricalcata sull'esperienza del Ministero degli affari esteri in cui tutti i documenti a partire dalla loro registrazione sono "tracciati" grazie a una registrazione efficiente (menzione di un oggetto, ma ugualmente di un luogo con un piano di classificazione, con un fascicolo), coprendo questo controllo tutte le tappe fino all'archiviazione del documento. L'altro gruppo di lavoro riguarda il miglioramento del modo di lavorare dei segretariati, di cui si sa il ruolo centrale in materia di controllo della produzione documentaria all'interno dell'ufficio: dalla registrazione della posta all'ordinamento e alla classificazione dei fascicoli, al versamento nel servizio d'archivio. Diventa più urgente allora la possibilità per l'archivista di elaborare in un modo

sistematico buone procedure in materia di trattamento dell'informazione, integrando le antiche pratiche basate sulla carta con i nuovi modi di lavorare indotti dalla generalizzazione dell'informatica negli uffici.

Per assicurare l'efficacia dell'archiviazione dei dati nel medio e nel lungo termine, la politica dell'organizzazione produttrice deve essere chiaramente stabilita e uno dei primi obiettivi da raggiungere è integrare la funzione archivi nel campo coperto dalle differenti applicazioni dell'istituto. Quando si sarà specificato il perimetro da coprire, cioè si sarà precisata l'esigenza di archiviazione, si stabilirà un'organizzazione con le relative persone, organizzazione che tiene conto dei legami tra queste persone e delle relazioni che esse alimentano.

Senza questa base, sarà molto difficile disporre dei mezzi finanziari e umani necessari per l'archiviazione dei dati. L'archivista deve essere associato all'elaborazione di questa politica per farsi intendere dagli informatici e per influenzarli. È importante non confondere le funzioni e far sì che la funzione archivio sia chiaramente identificata nei confronti di quella dell'informatico. Così l'archivista potrà svolgere il ruolo di mediatore neutrale tra l'informatico e ogni utente.

La conservazione a lungo termine dei dati digitali è un'operazione costosa. Ma ricostruire dati perduti, deteriorati e distrutti lo è ancora di più.

Questi costi importanti hanno un impatto sulla vita quotidiana dell'istituto di conservazione. Si tratterà dunque di fare delle valutazioni e delle selezioni prima di tutto in funzione dell'interesse scientifico dei dati, poi in funzione delle capacità umane e finanziarie dell'istituto di trattare questa prima selezione.

Non si cercherà di far prova dello stesso rigore per documenti destinati a una conservazione a lungo termine e per documenti eliminabili a breve scadenza (prima di dieci anni). Tuttavia essendo sempre più spesso l'archiviazione immediata un mezzo dei servizi d'archivio per giustificare la loro esistenza e svilupparsi, la questione dovrà essere esaminata in funzione delle situazioni. Del resto non si sottolineerà mai abbastanza l'importanza per l'archivista di sapersi posizionare molto a monte nei processi di creazione documentaria e di intervenire a partire dall'età di vita dei documenti.

La metodologia della preservazione che segue è dunque essenzialmente concepita per i documenti che, dalla loro creazione o dal loro trasferimento in un servizio d'archivio, sono destinati a essere conservati definitivamente.

I metodi di descrizione, il loro approfondimento, le scelte dei supporti dovranno essere più o meno diretti in funzione dell'interesse e della durata della vita dei documenti e delle esigenze di disponibilità e di comunicazione dei dati.

Bisogna allora che l'archivista esamini, insieme con l'istituto produttore, come razionalizzare la politica e la produzione documentaria e come renderla chiara a ogni utente.

Sempre più spesso, l'archivista si confronta con il desiderio dei gestori di mettere in opera dei sistemi globali di gestione elettronica di documenti (GED). Quest'orientamento nasce spesso dall'idea falsa che la dematerializzazione farà guadagnare preziosi metri quadrati che non saranno più dedicati a stoccaggio di documenti di carta. Se il discorso dell'archivista deve far sì che i decisori prendano coscienza delle riorganizzazioni interne importanti che implica una tale impresa per la produzione documentaria e deve far moderare un po' il loro entusiasmo, deve anche accompagnare questa pratica innovativa e ineluttabile. Un sistema di GED è un insieme composto di documenti che provengono da un gran numero di produttori differenti. L'archivista dovrà esaminare ognuno dei sottoinsiemi, fondandosi sul principio del rispetto dei fondi e dovrà esaminare ognuna delle tipologie all'interno dei suoi sottofondi. Ognuna avrà il suo metodo di trattamento appropriato, essendo certi documenti di tipo posta elettronica, altri di tipo immagini e documenti digitalizzati, altri di tipo fascicoli divisi, altri di tipo basi di dati.

L'archivista deve aiutare il produttore a esplicitare le sue intenzioni e gli obiettivi della produzione dei dati. Tutti questi elementi serviranno a alimentare la documentazione sull'oggetto d'informazione digitale da una parte, la riflessione sulla valutazione e la selezione di questi oggetti dall'altra. L'archivista è oggi uno dei personaggi chiave del *Record Management*, contribuendo a una formalizzazione dei sistemi di gestione e di produzione documentaria che non può essere che vantaggiosa per l'istituto.

IV. Preparare il versamento

Se l'istituto decide di avere una politica di *Record Management*, l'archivista è associato alla creazione dei documenti e le funzionalità di valutazione, di selezione e di comunicazione dei dati sono integrate nella copertura funzionale delle applicazioni.

Ma se, come è ancora frequente, l'archivista non è stato associato alla creazione del documento elettronico non più che alla gestione del suo ciclo di

vita, si trova spesso, come nel caso della carta, a confrontarsi con un'archiviazione di dati di cui non ha alcuna conoscenza preliminare. Cosa fare in questo caso?

IV.1. Reperire i documenti digitali esistenti

L'archivista prima di tutto dovrà provare a prevenire i versamenti anarchici, tentando di fare la rilevazione più completa della produzione dell'istituto.

Dovrà reperire le applicazioni elettroniche esistenti in modo che le informazioni non costituiscano più l'oggetto di una stampa cartacea, che sono le più fragili fra tutte.

Pertanto, chiederà prima di tutto ai servizi informatici del produttore la lista delle applicazioni di cui essi hanno conoscenza.

È così spesso obbligato a completare questa lista con i produttori stessi che possono aver prodotto delle applicazioni o delle schede diverse senza aver fatto appello ai tecnici.

L'articolo 22 della legge 78-17 del 6 gennaio 1978, riguardante la creazione della Commission nationale de l'informatique et des libertés (CNIL), precisa che "la commissione mette a disposizione del pubblico la lista dei trattamenti, che stabilisce per ciascuno di essi: la legge o l'atto regolamentare che decide della sua creazione o la data della sua dichiarazione; la sua denominazione e la sua finalità; il servizio presso il quale è esercitato il diritto; le categorie d'informazione nominative registrate così come i destinatari o categorie di destinatari abilitati a ricevere comunicazione di queste informazioni". Sono tenuti a disposizione del pubblico, nelle condizioni stabilite dal decreto, le decisioni, avvisi o raccomandazioni della commissione la cui conoscenza è utile all'applicazione o all'interpretazione della presente legge.

Per conoscere le applicazioni nominative, una fonte utile è dunque l'insieme degli atti regolamentari che riguardano decisioni di trattamento automatizzato d'informazioni nominative dopo avviso motivato della CNIL, per i trattamenti operati per conto dello Stato, di un istituto pubblico o di una collettività territoriale o di una persona morale di diritto privato che gestisce un servizio pubblico, così come i decreti del Ministero delle finanze riguardanti approvazione dei programmi d'inchiesta statistica obbligatori e d'interesse generale.

Finalità, nomi dei finanziatori e utilizzatori delle applicazioni, elementi di documentazione minimale delle applicazioni, potranno servire alla descrizione archivistica di quelle.

Va notato che certi atti regolamentari relativi a trattamenti riguardanti la sicurezza dello Stato, la difesa e la sicurezza pubblica non sono pubblicati. Si comprende così come certi programmi abbiano potuto non essere dichiarati presso la CNIL. In certe istituzioni, per ignoranza generale, questo caso può essere frequente.

Si ricorrerà infine a tutti i documenti delle direzioni delle finanze, delle conferenze di bilancio, nelle quali sono segnalate le applicazioni da mettere in opera nel prossimo esercizio. Quando esistono schemi direttivi informatici, si consulteranno i risultati di questi schemi allo scopo di giudicare realizzazioni e applicazioni messe in funzione. Questi bilanci danno spesso luogo alla programmazione di un nuovo schema che l'archivista dovrà attentamente prendere in considerazione per reperire i progetti futuri e tenersi informato dei loro sviluppi.

Esempio di archivio comunale

L'archivista del comune X chiede ai suoi colleghi dei servizi informatici di fare questo stato delle applicazioni. Riceve il quadro seguente (estratto da un quadro originale di 35 applicazioni per un comune di 45 mila abitanti):

Contenuto	Tipo d'informazione	Sistema	"Perte écrit"	Osservazioni
Gestione dell'aiuto sociale	Testo	Oracle		
	Testo	Oracle		
Gestione degli alunni	Testo	Access		
Informazioni catastali	Testo, immagini	Fox		
Permessi, ...				
Liste degli elettori	Testo	Oracle		
Atti di nascita, atti di decesso Matrimonio ...	Testo	Oracle		
Iscritti ai concorsi + note + risultati	Testo	Quattro		"dev interne"
Piante, catasti, informazioni catastali	Testo + scheda di formato "star"	Oracle + Star	Sì Piante	

Alcune osservazioni sul quadro precedente

Si tratta di una prima individuazione sulla quale gli archivisti e gli informatici potranno lavorare. Sarà necessario completare un tale lavoro sui seguenti punti:

- *formato preciso delle schede con l'identificazione delle loro versioni;*
- *versioni dei sistemi di gestione;*
- *basi aggiornate o no, periodicità degli aggiornamenti;*
- *date di creazione, di modifica e persino di conclusione;*
- *volume dei dati nel momento T;*
- *accrescimento annuale.*

Bisognerebbe ugualmente utilizzare dei termini più espliciti e archivistici. I titoli devono essere comprensibili da tutti gli utilizzatori. I servizi informatici del comune X hanno indicato sotto la forma "dev interne" il fatto che il programma sia stato scritto a loro cura e che essi controllano le informazioni.

L'espressione "perte écrit" non è più esplicita. Si tratta infatti delle basi che non danno luogo ad alcuna stampa cartacea. La nozione è fondamentale nel caso di un quadro di gestione, perché essa serve a determinare quale supporto sarà preferibilmente conservato rispetto ad un altro, ma essa deve essere espressa in altro modo per essere intelligibile da tutti.

Il servizio chiamato utilizzatore è in effetti il servizio produttore dei dati, cioè per gli informatici quello "che utilizza" la base. Sarebbe dunque bene tener conto anche del o dei servizi soltanto utilizzatori e non produttori dei dati, dunque dei destinatari dell'informazione. Inoltre nel quadro di un sistema di gestione elettronica di documenti condiviso e di sistemi intranet, i servizi produttori dei dati possono essere multipli. Bisognerà identificarli chiaramente e descrivere nei dizionari dei dati la parte degli uni e quella degli altri nella creazione.

Accade spesso che i servizi informatici non siano in possesso dei codici fonte delle basi utilizzate dai servizi produttori quando essi li hanno subappaltati. Questa carenza può intralciare la migrazione e la riutilizzazione dei dati in altri contesti tecnologici. L'ideale sarebbe che i servizi d'archivio fossero depositari di questi codici fonte allo scopo di perpetuare i dati. I codici fonte possono in effetti essere considerati come facenti parte dei metadati associati ai dati digitali. Gli archivisti dovranno sensibilizzare i servizi informatici a questo problema, soprattutto nel momento della scelta in materia di sviluppo di un'applicazione.

IV.2. Costituire i metadati

L'identificazione delle applicazioni non è che una tappa perché bisogna anche per ciascuna di esse raccogliere una documentazione molto precisa, costituita dai metadati (dati sui dati) del documento elettronico. I metadati danno informazioni sul contesto di creazione, di vita e di trattamento del documento elettronico. I criteri di valutazione, di selezione (tipo, data prevista, data di esecuzione ...) e di comunicazione possono anche, fin dalla creazione dei dati,

far parte dei loro metadati. Si hanno dunque differenti tipi di metadati: di descrizione, di creazione e gestione, di conservazione a lungo termine.

Esistono oggi numerosi standard di metadati. Il più conosciuto è quello che serve a descrivere sommariamente le pagine web, l'iniziativa del Dublin Core, che contiene 15 elementi. Per la conservazione a lungo termine, si potrà tenere particolarmente conto dello standard sviluppato dagli Archivi nazionali d'Australia o di quello sostenuto più recentemente dalla Library of Congress per le biblioteche digitali. Elementi come quelli della norma di descrizione ISAD (G), formati bibliografici MARC e la DTD EAD (Encoded Archival Description), strumenti conosciuti dagli archivisti e dai bibliotecari, possono essere ugualmente considerati come dei metadati. Anche la norma *Record Management* ne contiene e un sottogruppo di lavoro del TC 46/SC 11 dell'ISO la sta completando.

Certi metadati da raccogliere sono comuni a tutti i documenti digitali (base di dati, trattamento di testo, documento strutturato, immagine ...): sono essenzialmente quelli che servono alla descrizione dei documenti. Ma certi altri metadati esistono solo per certi tipi di documenti. Per una base di dati relazionale per esempio, sarà necessario recuperare in questa documentazione il dizionario dei dati, che non esiste per un documento burocratico. Per quest'ultimo, bisognerà ricercare le specificazioni di formato che non sempre sono pubbliche.

Questi metadati sono, per una parte di essi, lunghi da costituire e a volte difficili da raccogliere quando lo si fa a posteriori al momento dell'età intermedia o al momento del versamento in un servizio d'archivio definitivo, poiché i creatori e i produttori di documenti elettronici spesso non dispongono del tempo necessario per formalizzare la loro conoscenza della base che si fonda in gran parte sulla memoria. È indispensabile quindi disporre di tutta la documentazione associata ai dati. Se ciò costa un po' di tempo agli interlocutori degli archivisti, essi potranno tuttavia ricavarne un beneficio a termine quando avranno bisogno di ritrovare la traccia di dati elettronici o la ragione della loro assenza.

Un archivista può essere sicuro di aver raccolto i metadati indispensabili quando al momento della lettura dei dati elettronici che deve ricevere comprende perfettamente e in modo evidente il loro contesto di creazione e il loro contenuto. Deve pertanto mettersi al posto del ricercatore.

I metadati dovranno essere in futuro preferibilmente consegnati in formato elettronico. Ma essi possono ancora trovarsi esclusivamente su supporto cartaceo, soprattutto quando la forma elettronica non è stata preservata (documentazione associata a trattamento di testo per esempio). Fino a oggi, la

documentazione delle schede strutturate archiviate al Centro degli archivi contemporanei è così quasi esclusivamente in formato cartaceo. Quando i metadati sono in formato elettronico, essi possono essere incorporati nel documento stesso o costituire una o più applicazioni di documentazione separate dal documento. I metadati possono essere affinati in permanenza, completati o corretti se necessario. L'essenziale è raccoglierne il massimo per assicurare la sopravvivenza delle applicazioni. Oggi il documento elettronico può documentarsi interamente da sé stesso. Il miglior linguaggio per stabilire questa documentazione è l'XML.

Esempi di metadati

Metadati di conservazione indispensabili e minimali:

- *nome o identificativo (eventualmente codificato) del o dei file o dell'applicazione;*
- *dizionario dei dati per una base di dati;*
- *dizionario dei tag o degli elementi per un documento strutturato in SGML o XML, con la sua DTD o il suo schema;*
- *struttura dei file;*
- *struttura ad albero elettronica dei file e dei dati;*
- *codificazione dei dati per una base di dati (con la data di validità di questi codici);*
- *codici-fonte;*
- *sistema di gestione;*
- *software di produzione e versione.*

Metadati di descrizione:

- *titolo o intitolazione del documento;*
- *descrizione breve del contenuto e del contesto di produzione;*
- *servizio produttore (da dettagliare con gli elementi della norma ISAAR CPF);*
- *autore (i);*
- *mittente;*
- *destinatario (i);*
- *data di creazione, di modificazione (i), di cattura;*
- *formato elettronico;*
- *linguaggio dei dati;*
- *software di cattura e versione;*
- *supporto elettronico di origine;*
- *indicizzazione per materia;*
- *situazione nel piano di classificazione dell'istituto.*

Metadati di gestione:

- *servizio versante;*
- *data di trasmissione dei dati;*
- *indirizzo di localizzazione;*
- *supporto elettronico di conservazione;*
- *volumetria;*
- *diritti di riproduzione;*
- *comunicabilità;*

- *durata di conservazione.*

Si farà attenzione a raccogliere ogni altra informazione che serva a comprendere il processo e il contesto di creazione e di trattamento dei documenti e dei dati.

IV.3. Redigere il piano di classificazione del soggetto produttore

Alla base di una buona archiviazione è l'organizzazione della classificazione dei documenti il più presto possibile, e soprattutto al momento della loro creazione. Pertanto bisogna fare un'analisi funzionale dell'organismo produttore, preferibilmente al più alto livello.

“Il metodo analitico che mira all'elaborazione di una classificazione delle attività riflette tutte le attività del soggetto e le situa nel quadro definito dalle missioni e degli obiettivi che assicura direttamente o per delega”.

Da quest'analisi, elaborata dai produttori aiutati dall'archivista, deriva il piano di classificazione che è il riflesso della struttura funzionale del soggetto produttore.

Va notato che questa analisi funzionale e questo piano sono molto differenti dall'organigramma dell'istituto, strumento di gestione molto effimero, mentre le grandi funzioni sono più durature.

È necessario:

- strutturare questo piano di classificazione e questa struttura potrà ispirare il piano di denominazione delle schede;
- nominare o rinominare i file in modo intelligibile da tutti, produttori e ricercatori futuri;
- organizzarli logicamente gli uni in rapporto agli altri in questo piano. Si tratta dunque di formulare un modello a posteriori per l'informazione, se ciò non è stato fatto a partire dalla creazione dei dati.

In un piano di classificazione gerarchico:

- il primo livello riflette abitualmente la funzione esercitata dal produttore;
- il secondo livello descrive le attività costitutive della funzione;
- il terzo livello e i seguenti dettagliano le azioni o insiemi di azioni che compongono ogni attività.

IV.4. Redigere i quadri di gestione

Un quadro di gestione elenca le grandi categorie di documenti prodotti e ricevuti da una persona morale o fisica nell'esercizio delle sue attribuzioni. I quadri di gestione sono realizzati dai servizi produttori e dai loro archivisti. Per gli archivi pubblici, essi sono convalidati dalla Direzione degli Archivi di Francia e, in seguito, nessuna eliminazione può essere fatta senza il visto di quest'Amministrazione o di uno dei suoi rappresentanti.

Da una parte i dati elettronici devono essere considerati nei quadri che gestiscono la produzione cartacea. Ogni supporto (cartaceo, analogico, digitale) deve essere messo in relazione con l'altro. Servizi dipartimentali hanno cominciato, come la Direzione, a riflettere sulla questione: così stanno per apparire nuove forme di quadri di gestione.

Da una parte i quadri di gestione di nuova generazione devono fondarsi sui piani di classificazione gerarchici della produzione documentaria di un istituto. È necessario passare da una tipologia lineare, non strutturata, presentata nei quadri attuali, a una relazione fra i tipi documentari che riflettono la formulazione di un modello dell'informazione, per esempio in una base di dati relazionale. In questo modo, i criteri di selezione diventano più espliciti.

Per ogni categoria di documento, dovranno essere indicati:

- la stesura;
- la posizione nel piano di classificazione;
- una breve descrizione del suo contesto di produzione con le date di inizio di produzione;
- il suo peso in termini informativi e legali (valore primario, secondario, documento autentico ...);
- la durata di conservazione nel servizio;
- la scadenza di utilità amministrativa;
- la sorte finale del documento (eliminazione, conservazione, tipo di selezione);
- le relazioni con altre categorie o tipi documentari.

IV.5. Selezionare i dati

Avendo effettuato l'identificazione dei dati, l'archivista dovrà mettere in relazione i diversi supporti di produzione (cartaceo/elettronico) allo scopo di valutare il valore amministrativo e storico di ognuno.

Ogni conservazione a lungo termine ha degli obiettivi scientifici e di ricerca storica. Archiviare una selezione di dati digitali significa, come per i documenti su carta, assicurare per le generazioni future la preservazione della memoria di ciò che è stato un istituto, il suo contesto e il suo ambiente

nazionale e internazionale. La selezione dovrà essere periodicamente rivista in funzione di questi obiettivi a lungo termine.

L'archivista deve definire, insieme con il soggetto produttore, i dati che saranno un giorno da archiviare. Questi dovranno avere, al momento dell'archiviazione, una struttura determinata e convalidata.

Certe applicazioni elettroniche del tipo base di dati, concepite principalmente per ritrovare i fascicoli di carta ai quali si riferiscono le registrazioni, possono avere una ricchezza informativa molto grande. L'archivista deve vigilare sul contenuto sempre più sofisticato e complesso di questi documenti e degli usi che potrebbero esserne fatti per la ricerca storica in futuro.

La ricchezza dei nuovi tipi di documenti: l'esempio di una base di dati del Ministero della giustizia

L'applicazione GIDE gestisce il controllo dei detenuti negli istituti penitenziari a partire dalla loro immatricolazione, il controllo delle loro domande presso il giudice di applicazione delle pene, le loro attività (lavoro, attività socio-educative), la costituzione del loro gruzzolo, le visite ricevute, la gestione dei prodotti consumati in mensa, la disciplina, i trasferimenti ... Si tratta dunque di una base di conoscenza eccezionale sulla detenzione (più di 600 tavole di informazioni) e il fascicolo cartaceo non ha quasi più valore aggiunto in rapporto a questa base, se non la conservazione dei documenti prodotti all'esterno dell'istituto (che non si digitalizzano). A partire da questo momento, la scelta dei dati da perpetuare non ha più molto a che vedere con gli elementi stretti che danno una chiave d'accesso ai fascicoli cartacei.

Si possono selezionare dei dati all'interno di un'applicazione perché non si devono necessariamente conservare tutti i dati: in effetti essi possono risultare ridondanti da un'applicazione ad un'altra o una parte di essi può essere inutile nel lungo termine.

Si può anche scegliere di conservare una parte dei dati di un'applicazione in formato cartaceo e non obbligatoriamente in formato digitale. Si deciderà in funzione della frequenza e del modo di comunicazione desiderato, in funzione anche di criteri di rendimento finanziario.

Si ha interesse a rivedere per ogni tipo di documento i criteri di selezione che erano stati definiti per la carta. Se le selezioni specifiche restano ammissibili e si applicano dunque a dei dati, le selezioni sistematiche (su una percentuale, su un criterio alfabetico o geografico) sono meno difendibili nel caso di documenti elettronici. Così, nell'affare delle "scomparsa dell'Yonne", i fascicoli ai quali la Giustizia ha dovuto ricorrere erano quelli di affari classificati senza seguito, che si sarebbero dovuti eliminare secondo il regolamento in vigore e che per caso erano stati conservati.

Queste selezioni sono state in genere previste per serie d'archivio molto voluminose che occupano molto spazio nei magazzini d'archivio, considerate senza interesse amministrativo a medio termine e senza interesse storico. La selezione stessa di un campione mostra che la serie documentaria presenta un interesse, ma che essa non può essere integralmente conservata per mancanza di posto. In modalità elettronica i volumi non hanno la stessa importanza per dei dati testuali. Bisogna dunque rivedere queste selezioni sistematiche e probabilmente accettare di conservare la totalità dei dati di uno stesso tipo per una serie documentaria.

Gli obblighi fissati dalla CNIL

Gli avvisi resi dalla CNIL contengono delle prescrizioni relative all'eliminazione dei dati nominativi in un lasso di tempo relativamente breve.

Tuttavia l'articolo 5 della legge n. 2000-321 del 12 aprile 2000¹⁴ ha modificato l'articolo 28 della legge n. 78-17. È ormai previsto "che al di là della durata necessaria alla realizzazione delle finalità per le quali sono state raccolte o trattate, le informazioni non possono essere conservate sotto una forma nominativa se non in vista del loro trattamento a fini storici, statistici o scientifici". Da allora dunque la conservazione a lungo termine dei dati nominativi da parte dei servizi d'archivio competenti è possibile.

Esempio dell'applicazione GIDE (gestione dei detenuti negli stabilimenti penitenziari)

La CNIL aveva previsto, in seguito alla dichiarazione dell'applicazione, una conservazione dei dati che durasse 15 mesi dopo l'immatricolazione, fondandosi questi dati, secondo quegli stessi che hanno avuto in carico il fascicolo, solo su una dichiarazione riguardante l'applicazione che ha rimpiazzato GIDE. I dati erano dunque soppressi, benché la circolare sugli archivi dell'amministrazione penitenziaria prevedesse ben prima del trasferimento agli archivi dipartimentali una conservazione dei dati cartacei che durasse 10 anni e una conservazione illimitata dei registri di immatricolazione. L'amministrazione penitenziaria, essendo ben cosciente che delle ricerche potrebbero essere necessarie nei fascicoli cartacei, accompagna queste distruzioni dei dati con edizioni cartacee multiple permettendo di ricostruire un fascicolo cartaceo! L'archivista del Ministero della giustizia ha ricordato l'esistenza della circolare allo scopo di far

¹⁴ La legge riguarda i diritti dei cittadini nelle loro relazioni con le amministrazioni.

modificare la scadenza dei 15 mesi e di farla corrispondere a una DUA¹⁵ e, in un secondo tempo, per far conoscere le disposizioni della legge del 12 aprile 2000. Riguardando l'impatto di questa legge e della sua applicazione reale, ci si orienterebbe per il Ministero della giustizia verso una dichiarazione generale alla CNIL, per l'insieme del ministero che miri a organizzare l'archiviazione dei dati una volta scaduta la DUA e non più la loro soppressione.

IV.6. Conservare i dati in formato cartaceo, digitale o analogico

Secondo le applicazioni bisognerà porsi la questione della conservazione in formato elettronico, in formato analogico o in formato tradizionale cartaceo. Si deve archiviare in formato elettronico? È sempre possibile o utile?

Il DLM-Forum raccomanda di conservare in formato elettronico ogni documento prodotto in questa forma. Ma è senza dubbio prematuro essere così sistematici e bisogna cercare di adattare la pratica alla realtà dell'istituto da cui si dipende e ai suoi mezzi finanziari e umani.

È irragionevole considerare la conservazione di grandi basi di dati in formato cartaceo, perché i modi di ricerca in voluminose pile di *listings*¹⁶ sono inefficaci.

Invece la posta elettronica e tutti i documenti burocratici individuali sono ancora spesso stampati e classificati in fascicoli, senza che ciò ponga il minimo problema. Se il modo di classificare questi fascicoli è corretto e risponde ai bisogni dell'istituto in materia di ricerca dell'informazione, non ci sarà forse bisogno di considerare la loro conservazione in formato elettronico.

Questa pratica va senza dubbio a cadere progressivamente in desuetudine e si avrà interesse a operare una conversione di questa archiviazione verso il modo elettronico a partire da quando ciò sembrerà realizzabile.

¹⁵ Con "durata di utilità amministrativa" si intende la durata legale o pratica durante la quale un documento è suscettibile di essere utilizzato dal servizio produttore o dal suo successore. Alla sua scadenza si applica la decisione riguardante il suo trattamento finale. In sostanza il documento non può essere distrutto durante questo periodo, che costituisce la sua durata minima di conservazione. La circolare del primo ministro del 2 novembre 2001 sulla gestione degli archivi nei servizi e negli istituti pubblici stabilisce che, in base alla legge 78-17 del 6 gennaio 1978, gli organi incaricati della gestione degli archivi intermedi sono tenuti a garantire che la progettazione dei trattamenti informatici dei dati compresi negli archivi elettronici ne permetta la conservazione nei termini fissati dalle istruzioni relative alla durata di utilità amministrativa, che i documenti digitali da versare abbiano caratteristiche di autenticità e che siano accompagnati dai metadati indispensabili all'ulteriore gestione dei dati.

¹⁶ Elenchi.

L'archiviazione dei documenti di uno stesso tipo di fascicolo su molti supporti non può essere che una soluzione transitoria, perché è complessa da gestire in modo razionale e sistematico.

Bisogna digitalizzare i documenti cartacei?

L'archivista si trova sempre più spesso di fronte alle questioni o ai progetti dei decisori del suo istituto che, per guadagnare spazio di stoccaggio, pensano brutalmente di dematerializzare l'insieme della produzione, ivi compresa quella in formato cartaceo, di scannerizzarla nella sua totalità e di integrarla in un sistema di gestione elettronico dei documenti.

La digitalizzazione è ugualmente seducente per la comunicazione dei documenti e in un istituto per la gestione di risorse condivise.

Tuttavia non bisogna dimenticare i problemi tecnici, organizzativi e finanziari che pongono tali operazioni.

Le risposte varieranno dunque in funzione:

- delle capacità finanziarie dell'istituto, poiché la digitalizzazione ha un costo in materiali, in formazione;
- dell'interesse per un formato digitale dei documenti in rapporto al loro formato cartaceo;
- della durata di conservazione legale dei documenti.

Si confronteranno i costi di conservazione cartacea e digitale. Il costo della digitalizzazione e del trattamento documentario che vi è necessariamente associato dipende dalla qualità del documento d'origine, dal suo formato, dai volumi, dal livello di trattamento desiderato (semplice identificazione o indicizzazione) e dal loro grado di automazione.

L'archivista deve così suggerire che non c'è alcun interesse a digitalizzare documenti eliminabili a corto o a medio termine se i modi di classificazione tradizionale cartacea sono provati e sufficienti per trovare rapidamente l'informazione ricercata.

In un buon numero di istituti non si avrà grande interesse per qualche anno a digitalizzare i fascicoli del personale, fascicoli di supporti molto compositi, ma che conterranno ancora per qualche tempo numerosi documenti cartacei forniti dall'interessato (fogli di domande di congedo malattia o regolari,

lettere di motivazione, procedure con l'amministrazione in attesa della loro dematerializzazione totale). In questo caso una classificazione alfabetica dei fascicoli come la praticano i segretariati abitualmente permette di ritrovare facilmente i dati ricercati e questa classificazione cartacea è ancora del tutto sufficiente. Sarebbe più oneroso e non molto più efficace per la ricerca considerare la digitalizzazione dei documenti cartacei forniti dagli individui e integrarli in un programma di gestione elettronico di documenti e in seguito gestirli sul lungo termine.

Se si decide di digitalizzare, bisognerà vigilare sulle scelte di risoluzione delle immagini, di formati. Si opterà per quelli che assicureranno al documento il maggior uso (la qualità d'immagine sia sufficiente perché non sia necessario ricorrere continuamente al documento cartaceo) e la maggior durata. Su questo punto si seguiranno le ultime raccomandazioni, in particolare la *Guide pour l'archivage des documents électroniques* prodotta dal MTIC. Ma queste raccomandazioni hanno anche delle conseguenze, di uguale peso sulle capacità di stoccaggio in linea o fuori linea.

IV.7. Garantire l'identità e l'integrità dei dati elettronici

L'autenticità poggia per il momento sul postulato che, al momento del versamento nel servizio di conservazione a lungo termine, il contenuto dei dati fosse autentico.

Il gruppo InterPares (International Research on Permanent Authentic Records in Electronic Systems) ha determinato un insieme di otto regole che garantiscono l'autenticità: 1) associazione di metadati di descrizione ai dati, 2) diritti d'accesso, 3) procedure di protezione contro la perdita e la corruzione dei dati, 4) procedure di protezione contro il deterioramento dei supporti e contro l'obsolescenza tecnologica, 5) istituzione di schede associate a ogni procedura, 6) regole di autenticazione dei dati, 7) identificazione dell'esemplare che fa fede in caso di esistenza di copie multiple, 8) documentazione che deve essere rimessa a chi assicura la conservazione a lungo termine quando i dati diventano storici. Più il numero di regole rispettate è vicino alle otto, più l'autenticità del documento ha possibilità di essere forte. Ma un documento elettronico che non rispetta alcuna di queste otto regole non sarà tuttavia distrutto. Bisogna far prova di spirito pragmatico in questo campo e non cadere in discussioni troppo teoriche. Non bisogna prendere più precauzioni sull'autenticità di certi documenti elettronici di quante non se ne siano prese per le loro forme su carta e in microfilm.

Il gruppo InterPares ha anche determinato quando le otto regole sono rispettate; un certificato di autenticità sarebbe fornito solo su espressa domanda.

Riflessioni sulla firma elettronica

Come su un documento cartaceo, la firma in formato elettronico è soltanto un anello di una catena complessa di fattori e di procedure che, insieme, garantiscono l'autenticità dei documenti.

Un decreto (n. 2001-272) relativo alla firma elettronica è stato emesso il 30 marzo 2001 per l'applicazione dell'articolo 1316-4 del codice civile. Esso determina le condizioni necessarie affinché procedimenti di firma elettronica possano essere considerati come sicuri e beneficiare della presunzione di affidabilità. Nell'amministrazione, si è ancora alla semplice sperimentazione, per esempio per le teledichiarazioni, per la trasmissione degli atti sottomessi al controllo di legalità. Ma a termine l'essenziale degli atti amministrativi è contenuto nel decreto elaborato per le scritture private.

Il decreto del 30 marzo permette teoricamente una larga accezione della firma elettronica (firma crittografata, firma biometrica, firma-tatuaggio, firma digitale). Tenuto conto di questo, esso precisa che "si presume l'affidabilità di un procedimento di firma elettronica fino a prova contraria quando questo procedimento mette in opera una firma elettronica sicura, stabilita grazie a un dispositivo sicuro di creazione di firma elettronica e [precisa] che la verifica di questa firma si fonda sull'utilizzo di un certificato elettronico qualificato". Non tutti i documenti disporranno evidentemente di questo tipo di firma.

La firma crittografica fa parte di quelle che possono rispondere alle esigenze di sicurezza definite nel decreto del 30 marzo 2001. Essa si fonda sull'esistenza, per il mittente e per il destinatario di un documento, di una chiave pubblica e di una chiave privata che permettono di realizzare la firma, eventualmente la trascrizione in cifra e la lettura del documento firmato. La chiave pubblica di ciascun individuo è resa disponibile all'interno di un annuario, mentre la chiave privata è conservata segreta. Benché la chiave pubblica e la chiave privata siano complementari, è impossibile, anche conoscendo la chiave pubblica, dedurre la chiave privata.

Il certificato, rilasciato da un'autorità di certificazione, contiene delle informazioni che permettono di associare una chiave pubblica a un individuo e di assicurarsi dell'identità del mittente di un messaggio.

Questa firma tende ad assicurare la circolazione dei documenti sulle reti nella misura in cui il suo sistema permette di reperire ogni alterazione del documento. Invece, per il momento, non è previsto niente che permetta una verifica della firma, anni dopo la sua istituzione. Ricercatori europei lavorano a un protocollo che permetterebbe la conservazione del contesto della firma.

Altrimenti è impossibile assicurare la perennità di una firma crittografica nella misura in cui, per assicurare la leggibilità del documento nel tempo, questo dovrà costituire l'oggetto di migrazioni. Il documento ne risulta inevitabilmente modificato, il che farebbe fallire i meccanismi di verifica, consistendo il principio della firma crittografica nel garantire l'identità e l'integrità del documento.

La *firma digitale* consiste semplicemente nella cattura, all'interno di un file informatico, dell'immagine della firma manoscritta di un individuo. L'immagine informatica risultante può successivamente essere aggiunta, in seguito o all'interno di un documento elettronico e conservata nello stesso formato.

Esempio: il sistema SAGA sviluppato per il Servizio centrale dello stato civile del Ministero degli affari esteri (SCEC) a Nantes

Il sistema comporta tre elementi distinti: prima di tutto, gli otto milioni di atti digitalizzati a partire dai registri cartacei; in seguito il sistema che permette all'ufficiale di stato civile un "pavé" contenente il sigillo dello Stato e la firma all'interno dell'atto digitale; finalmente, la carta assicurata su cui la copia conforme firmata è stampata, documento provvisto di caratteristiche speciali che proteggono la sua integrità e ne impediscono la riproduzione.

All'interno dello SCEC, la sicurezza è assicurata da un insieme di metodi e di tecniche: gli ufficiali di stato civile non sono un qualunque gruppo di utilizzatori e la penalizzazione severa del falso in scrittura pubblica assicura, più di quanto tutte le misure tecnologiche potrebbero fare, che essi non siano considerati come dei potenziali falsificatori all'interno del sistema. Meccanismi di aggiornamento automatico delle procedure informatiche assicurano la loro tracciabilità. L'accesso ai locali contenenti i documenti digitalizzati e al documento assicurato è controllato dalle procedure tradizionali di chiavi e serrature e l'accesso ai posti di lavoro degli ufficiali di stato civile è controllato da una password. Di più, l'utilizzazione della firma digitale dell'ufficiale è strettamente guidata dal sistema: l'immagine digitale della firma non si trova sul posto di lavoro dell'ufficiale, ma su un server centrale, essendo trasferita solo al momento dell'identificazione dell'ufficiale sul suo posto di lavoro; la firma non è utilizzabile se non nelle modalità definite dal sistema e non può essere estratta per un'utilizzazione non conforme.

Per dei documenti digitali utilizzabili a fini giuridici, la firma elettronica costituisce una delle garanzie essenziali della loro integrità e ogni alterazione

costituisce una minaccia per questa. Ma per gli altri documenti, un dubbio riguardo alla loro autenticità può essere tollerato e la firma avrà meno importanza e potrà essere separata dai dati se essa impedisce la loro conservazione tramite migrazione. Il ricercatore userà altri elementi di critica e metodi tradizionali della diplomatica per giudicare il degrado dell'integrità dei dati.

La forma materiale è anche fino ad ora un elemento determinante per giudicare l'integrità del documento. La tecnica correntemente più usata fino ad oggi consiste nel liberare i dati dal loro contesto di formattazione, per preservarli più comodamente nel lungo termine. Un elemento importante per giudicare l'integrità dei documenti è così perduto. Questa constatazione è piena di conseguenze per la dematerializzazione degli atti autentici, il cui formalismo è molto importante.

IV.8. Rendere sicuro il trasferimento

È necessario che il servizio che deve conservare i documenti disponga almeno di una stazione di lavoro con i lettori appropriati ai formati da verificare in consegna.

Importanza dell'attrezzatura tecnica: esempio dell'archiviazione della Nouvelle Chaîne Pénale di Parigi

La Direzione degli Archivi di Francia desidera accompagnare l'archiviazione di questa applicazione negli archivi dipartimentali di Parigi. La sua attrezzatura le permetterebbe di leggere dei CD o delle cassette DAT 4 mm. I servizi del Ministero della giustizia però non avevano masterizzatore di CD connesso al server Unix nel quale erano immagazzinati i dati e potevano consegnare soltanto delle cassette DAT 8 mm. La storia finisce bene perché la Direzione informatica della città di Parigi ha recuperato i dati della cassetta DAT 8 mm, li ha immagazzinati e poi masterizzati su un CD rimesso alla Direzione degli Archivi di Francia per i suoi test.

L'archiviazione dei dati digitali necessita di stretta collaborazione e di scambio di mezzi.

Il servizio d'archivio informerà in anticipo i servizi produttori che saranno indotti a fare dei versamenti, dei formati e dei supporti sui quali questi versamenti dovranno essere fatti.

Quando l'archivista non è stato associato a tutta la catena di trattamento del documento, egli esaminerà a monte molto accuratamente il contenuto dei versamenti prima del loro ingresso.

I file saranno aperti, eventualmente a campione se essi sono troppo numerosi. I loro metadati saranno controllati e la loro raccolta eventualmente proseguirà finché essi non saranno il più possibile completi. L'archivista si assicurerà così che ciò che egli integra nelle loro collezioni è conforme alla produzione iniziale e che i dati sono integri. Questo lavoro è molto lungo, il che giustifica pienamente un intervento precoce dell'archivista nella vita del documento: ciò eviterà pesantezze e perfino insuccessi alla fine della catena.

La registrazione dell'entrata sarà effettuata sul sistema usuale dell'istituto.

V. Garantire la conservazione

V.1. Condizioni minime richieste per l'accettazione di un versamento di dati elettronici

Come per i supporti tradizionali, la prima delle condizioni consiste naturalmente nell'assicurarsi che i dati e i documenti da accogliere siano quelli che ci si attendeva, verificando con un sondaggio e prima del versamento il contenuto della distinta di versamento con il contenuto del documento.

Ma il versamento ha delle caratteristiche nuove da analizzare. Bisognerà non soltanto verificare i contenuti, l'esistenza e la trasmissione di tutti i metadati indispensabili alla sopravvivenza dei documenti elettronici, ma anche il formato dei dati.

Si verificherà per esempio

- *l'esistenza e la trasmissione delle liste e delle definizioni (dizionario) dei dati, dei loro codici e delle loro strutture. Si verificherà che questa documentazione descriva bene il o i file da versare. Documentazione e dati digitali sono indissociabili e devono essere versati nello stesso momento;*
- *la conformità del formato immagine in rapporto al formato annunciato (es.: PNG, TIFF);*
- *la conformità di un documento XML, in rapporto alla DTD o allo Schema annunciato e il versamento di questa DTD o di questo Schema.*

V.2. Scelta dei formati e dei supporti

Ogni applicazione fuori linea si fonda su formati e su supporti che necessitano del materiale di lettura specifico. Per conservare i dati nel loro formato d'origine, bisognerebbe dunque costituire musei di tecnologie informatiche utopistici, tanto considerevole sarebbe il loro costo. È per questo che ci si è fino ad oggi dedicati a privilegiare la conservazione dei dati, dell'informazione e non quella della forma e del formato dei documenti d'origine.

V.2.1. Formati e linguaggi di codifica

All'interno di un sistema informatico, ogni informazione è necessariamente rappresentata alla base sotto forma di un codice binario composto da 0 e da 1 (i bit).

Numerosi formati di codifica permettono in seguito di presentare questa informazione in formato testuale, immagine Per esempio un formato testo è composto da caratteri che sono essi stessi composti da combinazioni di otto 0 e 1 (byte). È a questo livello di codificazione che si pratica in generale l'archiviazione a lungo termine dei dati.

Ma un formato di codificazione non è che un elemento in un insieme di strumenti che lo rende intelligibile: un file Word per esempio è sempre concepito per essere gemellato con una versione del programma Word, esso stesso concepito per un certo modello di computer e per il suo proprio sistema di utilizzo. Un documento elettronico non è leggibile se non grazie all'interazione della codificazione con un programma e del materiale informatico che permette di interpretarlo.

Ci sono dunque diversi livelli di codificazione che sono incastrati gli uni negli altri.

I formati di codificazione possono essere:

- aperti (se le specifiche sono pubbliche);
- chiusi (se le specifiche sono tenute segrete dal proprietario);
- proprietari, cioè definiti da un'impresa privata e sottomessi a dei diritti; un tale formato il cui uso può essere gratuito può diventare a pagamento domani se cambia la politica commerciale del suo proprietario; un formato proprietario può essere aperto;
- standard, cioè prodotti da un organismo di normalizzazione (ISO, AFNOR, W3C); sono aperti e non proprietari.

I formati testuali

La loro codificazione si fonda su un formato testo, al quale basta gestire i caratteri alfanumerici, come un certo numero di caratteri “bianchi”. Il più spesso, questi caratteri sono codificati in ASCII, ma questa codificazione, che non specifica la rappresentazione dei caratteri accentati, potrebbe essere sostituita dall’UNICODE. L’ASCII e l’UNICODE sono norme ISO. Questa normalizzazione facilita la conservazione dei dati digitali.

Un documento può essere creato in puro formato testo grazie ad un *editor* di testo, programma la cui funzione principale è di permettere l’edizione (creazione/aggiunta/modifica/soppressione) di testo semplice, senza alcun arricchimento tipografico o messa in forma. Ne risulta un file in TXT, che contiene questo testo grezzo.

Altrettanto quando per conservazione a lungo termine si vuol mettere un documento in un formato grezzo, in ASCII o in UNICODE, ogni formattazione originale scompare (font del carattere, misura, italico o grassetto, ecc.). Si conservano soltanto gli spazi e i caratteri.

Malgrado la normalizzazione di questi linguaggi, questo tipo di formattazione non è sempre facile nella misura in cui ciascuno dei sistemi di gestione (Windows, Macintosh, Unix, Linux ...) codifica il formato testo diversamente.

A partire da questi linguaggi, si sono sviluppati dei trattamenti di testo, dei meta-linguaggi e linguaggi di marcatura, come dei formati misti che combinano testo e immagine.

I trattamenti di testo

Un trattamento di testo è un programma che serve da editor del testo, in modo più potente, più completo ma anche più complesso di un semplice editor di testo. Uno dei più conosciuti è Word (programma proprietario e chiuso) concepito da Microsoft. Il formato testo messo in opera da un trattamento di testo, concepito per rispondere a bisogni di edizioni modeste, dà al testo lo stesso aspetto sullo schermo e su carta. Lo strumento può anche permettere di strutturare l’informazione in modo assai sofisticato, ma non normalizzato.

Un documento può essere realizzato grazie a una macro o macro-comando, comando formato da una successione di altri comandi, più o meno strutturati. I linguaggi di macro sono sempre più completi, in modo che una macro va dal semplice doppio-clic automatizzato a una piccola applicazione. Questi comandi sono molto specifici e legati alla versione di un programma. Essi

sono anche fonti di numerosi virus. È dunque imperativo liberare i file da questi comandi per la lunga conservazione.

L'impossibilità di scambio tra programmi di trattamento di testo, compreso quello fra le differenti versioni di uno stesso programma, ha condotto la Microsoft a produrre il formato RTF (*Rich Text Format*). È un formato proprietario, ma aperto. Specifica dei parametri relativi alla formattazione. Tutti i programmi commerciali si presume che lo leggano, per assicurare un minimo di interoperabilità con Microsoft Word, ma si sono constatate numerose anomalie e non-portabilità di file creati in questo formato.

È praticamente impossibile assicurare la perennità di questi formati specifici di trattamento di testo. Al contrario, i linguaggi di marcatura presentano delle garanzie.

I meta-linguaggi e i linguaggi di marcatura

All'origine dei linguaggi di marcatura e di strutturazione dell'informazione è l'SGML (Standard Generalized Markup Language), meta-linguaggio di marcatura normalizzato nel 1986, utilizzato dagli editori, dagli universitari e soprattutto dal mondo dell'industria per creare grandi documentazioni tecniche. Essendo questo meta-linguaggio pesante e di complessa utilizzazione, il W3C ha sviluppato l'HTML (Hyper Text Markup Language) per la creazione semplice e rapida di documenti "ipertesto". Non proprietario e aperto, esso fa parte degli standard, ma è utilizzato molto diversamente secondo gli editori di pagine Web.

L'HTML presenta degli inconvenienti: gli utilizzatori non possono definire una struttura che sia loro propria e questa struttura mescola la forma e il fondo, cioè a una struttura è obbligatoriamente associata una rappresentazione. Si presta male alla creazione di "iperdocumenti" complessi.

È per questo che il W3C ha deciso di produrre un insieme di raccomandazioni, l'XML (eXtensible Markup Language), la cui prima versione, XML versione 1.0, pubblicata nel 1998, definisce il meta-linguaggio di base. L'XML deriva dall'SGML, ma è adattato alle nuove esigenze del Web. La finalità di questo meta-linguaggio è lo scambio di dati e l'interoperabilità delle applicazioni. La combinazione di XML e del linguaggio di programmazione di formattazione XSL-T (eXtensible Style language-Transformations), può rispondere a tutti i bisogni di edizione, qualunque sia la piattaforma informatica, qualunque sia l'applicazione. Un documento XML può essere presentato in tanti modi, grazie a XSL-T.

In HTML, descrizione fisica e strutturata sono intimamente mescolate. Ciò conduce a due maggiori difficoltà:

- i documenti creati hanno una speranza di vita breve, limitata dalla perennità delle tecnologie capaci di dare agli elementi di messa in pagina l'interpretazione attesa;
- uno stesso documento non può essere facilmente realizzato secondo modelli fisici differenti, il che limita la sua accessibilità.

Invece il fatto che contenuto strutturato e presentazione dei dati siano distinti nell'XML può porre dei problemi per la critica dell'integrità dei documenti e della loro autenticità, come è stato sollevato nel rapporto sulla dematerializzazione degli atti autentici.

Si può, se lo si desidera, modellizzare tipi di documenti predefiniti in XML grazie a DTD (Document Type Definition) o a schemi XML. I dati non sono in questo caso interpretabili se non si conosce la DTD o lo schema che bisognerà dunque aver cura di conservare come referente nei metadati.

I formati immagine

La scannerizzazione è il modo più semplice di digitalizzare gli archivi esistenti su carta, ottenendo un'immagine elettronica di ogni pagina di un documento e collegando queste pagine con un meccanismo di indicizzazione appropriata. Il formato immagine non assicura nessuna struttura semantica ai dati contenuti nell'immagine. In particolare il testo scritto è compreso in quanto immagine e non in quanto sequenza di caratteri intelligibili. Affinché il computer sia in grado di effettuare trattamenti sul testo contenuto in un file immagine, bisogna effettuare un'ulteriore operazione di riconoscimento ottico dei caratteri.

L'immagine può essere presa con un apparecchio fotografico digitale o con uno scanner piatto. La sua codificazione si fa secondo una quadrettatura (costituita da punti luminosi o pixel più o meno grossi sullo schermo), la cui densità definisce la risoluzione dell'immagine. Per la rappresentazione informatica, si dà a ogni *pixel* un valore codificato su un certo numero di bit, che va a determinarne il colore o l'intensità. Il numero di *pixel* per pollice detto "dpi" (dot per inch) (100, 300, 600...dpi) sarà scelto in funzione della qualità richiesta per l'immagine digitale. Maggiore è il numero di punti su una superficie data, migliore è la qualità dell'immagine, ma più aumenta il suo volume in bit. La compressione permette di ridurre questo volume, ma certi algoritmi di compressione degradano l'immagine d'origine. Più un'immagine

è degradata dunque, meno essa sarà perenne, poiché ogni migrazione dei dati e su supporti rischia di farle perdere ulteriormente la sua qualità. Nel caso in cui l'algoritmo di compressione non faccia degradare l'immagine, quella ha ogni possibilità di evolvere meglio nel tempo. Si farà attenzione anche ad avere, per il passaggio dalla generazione n alla generazione n+1, lo stesso strumento di compressione/decompressione, perché in questo caso non si fa che copiare i file senza far subire loro alcuna trasformazione, il che è più rassicurante.

Esistono numerosi formati immagine. Si opterà per l'uno o per l'altro in funzione delle finalità della digitalizzazione dei documenti. Si sceglierà in funzione del *display* e della leggibilità desiderati, del posto occupato sui supporti. Bisogna ricercare il miglior compromesso tra il volume dei documenti elettronici prodotti e la qualità ricercata in funzione degli usi previsti.

Per la conservazione a lungo termine, bisogna privilegiare un formato come il TIFF (Tagged Image File Format) che supporta una compressione senza perdita; l'inconveniente è che l'immagine TIFF è voluminosa. Altri formati possono essere in concorrenza con questo: il PNG (Portable Network Graphics), che, compresso, è molto poco degradato e poco voluminoso e la nuova generazione di JPEG (JPEG 2000), ma facendo attenzione a utilizzare i parametri buoni.

I formati misti (testo e immagine)

L'utilizzazione di questi formati impone di scegliere a priori e irrimediabilmente la risoluzione del file che si crea. Il linguaggio PostScript propone una soluzione per risolvere questo problema, descrivendo delle immagini in un modo uniforme qualunque sia la risoluzione della periferica di edizione, con l'aiuto di equazioni vettoriali. Un documento scritto in linguaggio PostScript può essere inviato a ogni periferica che supporta il linguaggio PostScript, qualunque sia la sua risoluzione, producendo un risultato adatto a ogni tipo di periferica.

Il formato PDF (Portable Document Format) è basato sul linguaggio Post Script e riprende la sua filosofia. Quando l'obiettivo del PostScript è di poter rappresentare una pagina in modo che sia indipendente dalla risoluzione della periferica di edizione, l'obiettivo del PDF è di poter ripresentare un documento nello stesso modo, indipendentemente dalla piattaforma e dal computer utilizzato. Non si scrive direttamente un documento in PDF, ma si produce il documento nel programma appropriato e in seguito si trasforma il documento in PDF, più spesso grazie al programma Acrobat della società

Adobe. Il formato incorpora degli aspetti interattivi: legami ipertestuali, segnalibri, firma elettronica, cifratura ecc. Il formato PDF tenta dunque di realizzare tutto il potenziale del documento elettronico. È portabile in un gran numero di ambienti e la politica praticata dalla società Adobe, proprietaria del formato, di distribuzione gratuita del programma per la lettura dei documenti PDF ha contribuito a questa universalità. La conversione in PDF implica invece l'utilizzo del programma Acrobat Capture che è a pagamento.

Questi due formati sono proprietari ma aperti perché le specifiche sono conosciute. Sono dunque meno dipendenti di altri da programmi particolari. Il PDF può dunque essere eventualmente accettato per una conservazione a lungo termine.

Criteri di scelta dei formati

Per la conservazione a lungo termine, in funzione dei bisogni e delle capacità dell'istituto:

- si sceglieranno dunque preferibilmente linguaggi e formati standard o che costituiscono oggetto di raccomandazioni internazionali (XML, HTML, TXT per i documenti testuali).

In caso di impossibilità, si accetteranno dei formati aperti (PDF per i documenti testuali, TIFF o PNG per le immagini) e si eviteranno i formati chiusi.

V.2.2. Supporti

La conservazione a lungo termine è anche una questione di affidabilità dei supporti e di perennità dei materiali. Essa implica una revisione periodica delle scelte operate allo scopo di verificare la conformità delle soluzioni e delle soluzioni adottate con le evoluzioni della tecnologia, perché l'esperienza prova che ognuno dei due elementi della coppia supporto-materiale può rivelarsi obsoleto entro alcuni anni.

I criteri di scelta dei supporti sono:

- Il livello d'integrità richiesto per i dati.

La miglior garanzia in questo campo è, secondo la norma AF Z 42-013, il ricorso alla tecnologia del disco ottico WORM, che impedisce ogni modifica dei dati registrati.

- I mezzi finanziari disponibili.

Il livello di questi mezzi condizionerà direttamente il ritmo di rinnovamento dei supporti; il minimo indispensabile (cioè il livello al di sotto del quale è illusorio pretendere di fare l'archiviazione elettronica) è ciò che permette di effettuare l'operazione almeno due anni prima della data di scadenza dei supporti.

L'ideale sarebbe tuttavia essere in grado di procedere a rinnovamenti molto più precoci di quanto implichi la durata di vita annunciata dei supporti o della loro tecnica: perché più lungo è l'intervallo tra due migrazioni, più è importante il rischio di perdita della qualità dei dati (una migrazione che interviene alla fine di otto anni provoca una perdita di circa il 3%, mentre la perdita è ridotta allo 0,2% se la migrazione interviene nei due anni).

- Il volume dei dati da conservare sul lungo termine e la loro frequenza di utilizzazione.

Questo criterio è determinante per le scelte dei supporti. Dati molto voluminosi potranno trovar posto su cassette, di tipo DLT, 3480 o 3481 IBM. Ma questo supporto non è utilizzato dal grande pubblico e non si presta molto dunque alla comunicazione dei dati consultati spesso (per esempio basi di dati che servono come strumenti di ricerca). La misura dei supporti deve all'inizio essere adattata a quella dei file: un solo file conservato su un supporto che gli è ben adatto quanto al volume, è più rapidamente reperito da un sistema di gestione automatizzato, rispetto a numerosi piccoli file conservati su un supporto di forte densità.

- Al contrario, la longevità dei supporti non può essere oggi un criterio di scelta.

La maggior parte di essi degradano anche in condizioni di conservazione ottimali, entro 10 o 20 anni. Si stima che solo il disco di vetro non si deteriora e potrebbe essere ancora letto con i materiali esistenti per una cinquantina d'anni. Invece niente garantisce che i materiali di lettura supporteranno ancora per qualche decennio il loro formato. Perché un supporto di buona qualità in condizioni ottimali di conservazione può essere conservato senza alterazione, mentre la sua tecnica e i suoi strumenti di lettura saranno divenuti obsoleti e difficili da trovare sul mercato.

Esempio: i file immagine

Si è visto che è consigliato archiviare i file immagine piuttosto in formato PNG o in formato TIFF.

Un'immagine TIFF occupa, secondo la sua compressione, da 10 a 500 Megabyte. Un CD-R ha una capienza massima di 650 Megabyte. Si immagazzinano dunque su un CD-R soltanto da una a 60 immagini al massimo, il che è poco per grandi serie documentarie.

Non si archiveranno dunque grosse quantità di immagini su questo supporto; si masterizzeranno dunque soltanto i file richiesti molto frequentemente in consultazione.

Il Centro degli archivi contemporanei ha anche fatto la scelta di cassette SuperDLT per l'archiviazione in modalità immagine del censimento della popolazione 1999, essendo determinato il volume dei dati (14 terabyte) e il costo della loro archiviazione.

Invece, alcuni stimano che più la capacità del supporto di conservazione è importante, più lo è anche il rischio che corrono i dati.

Bisogna tuttavia tener conto di altri parametri per valutare i rischi: duplicazione dei dati, su supporti fisici differenti, immagazzinati in luoghi differenti, esistenza di procedure per sorvegliare lo stato dei supporti.

L'ideale sarebbe considerare la conservazione su due di questi tipi di supporti, in funzione dell'obiettivo di conservazione:

- per documenti molto richiesti dal pubblico in linea o su ordinazione (acquisto, prestito), si utilizzerebbe il disco. Ma in questo caso, bisogna assumere uno stoccaggio più pesante da gestire (su scaffali o su torri di CD);
- per documenti poco consultati e voluminosi, si utilizzerebbero delle cassette come le DLT;
- si può ugualmente combinare un supporto digitale e un supporto analogico.

Prima di utilizzare un qualunque supporto di conservazione, si potrà verificare o farsi assicurare dal fornitore che questo supporto e il suo lotto non comportano difetto di fabbricazione non rivelabile per l'utilizzatore. La durata di vita dei dati sarebbe molto ridotta. Il difetto può non essere constatato che al momento in cui si cercherà di aver accesso ai dati archiviati. Si farà attenzione dunque, dopo aver copiato dati sul supporto, alla qualità di lettura di quelli. Si integrerà peraltro, in un programma di gestione che può essere automatizzato, la verifica regolare dei supporti e si assicurerà loro attenzione tecnologica per essere pronti a cambiarne e a farne migrare i dati. È indispensabile che il BLER (Bloc Error Rate), indicatore di qualità, sia tale che i dati siano in classe 1 (o al peggio 2), perché la loro integrità sia garantita.

Certi esperti stimano che il solo modo di conservare realmente nel lungo termine è di conservare una copia in linea, che sarà aggiornata regolarmente verso nuovi supporti e alla quale si potrà ricorrere ogni volta che un problema subentrerà al supporto fuori linea e viceversa. Se le domande di comunicazione sono costanti si avrà interesse a lasciare i documenti in linea su un server.

Molto spesso, per varie ragioni (per preservare la confidenzialità di certe informazioni contenute o modificate in campi di una base o perché queste informazioni devono essere decodificate o modificate per essere rese comprensibili al pubblico, o per ridurre i tempi di telecaricamento, o per evitare il saccheggio ...), si mettono in linea file differenti. In questo caso può anche essere interessante, perfino necessario, archiviare fuori linea la versione messa in linea.

È in ogni modo indispensabile nel quadro di una buona archiviazione di documenti elettronici, prevedere molti esemplari degli stessi documenti, se possibile su molti tipi di supporti differenti. È la procedura della “ridondanza”. Più si trovano per esempio dei server che conservano esemplari dello stesso documento digitale, meno si corre il rischio di vederli distrutti tutti nello stesso tempo. Nel mondo digitale, la condivisione fa parte della preservazione a lungo termine. Si farà attenzione anche a non privilegiare mai un supporto per facilitare soltanto la comunicazione. Il principio di conservazione a lungo termine resta il criterio più importante.

Esempio: il Servizio di trasferimento e di archiviazione dei file (STAF) del Centro nazionale di studi spaziali (CNES)

Il CNES sceglie i suoi supporti di conservazione in funzione delle richieste di consultazione e di aggiornamento e in funzione dei bisogni di perpetuità dell'informazione: un dato molto richiesto sarà su dischi magnetici in linea, un dato meno richiesto sarà “near-line” con immagazzinamento delle cassette in una torre di cassette. Invece nessun documento è per i ricercatori del CNES fuori linea, disposto su palchetti. I produttori di dati del CNES determinano il livello di prestazione richiesto in materia di restituzione dei dati archiviati e di garanzia della loro perpetuità. Il supporto considerato dallo STAF come garante della migliore prestazione è il disco magnetico. I dischi sono automaticamente duplicati. A ciascun disco, corrisponde un disco specchio (che risulta dalla suddivisione di ciascun disco). Il supporto di secondo livello di prestazione è la cartuccia delle scale Storagetek o IBM. Le cartucce sono immagazzinate nelle librerie automatizzate. Lo STAF gestisce automaticamente la migrazione dei dati dai dischi verso le cartucce, ma il ritorno delle cartucce sui dischi scatta sistematicamente quando l'utilizzatore vuole una restituzione dei dati. Il sistema gestisce ugualmente il riciclaggio dei supporti: un supporto deteriorato è reperito dal sistema che lo ricicla. Le cartucce sono anche riciclate sistematicamente ogni cinque anni con lo stesso procedimento.

È così che per l'immagazzinamento, lo STAF propone:

- *supporti in linea (i dischi magnetici);*
- *supporti “quasi in linea” (near-line), le cartucce: la scadenza di restituzione è un po’ più lunga.*

La distruzione dei supporti originali che contengono i file può essere decisa soltanto molto dopo la migrazione di questi, perfino non completamente. Invece se si decide di conservare questi supporti originali, si farà attenzione a conservare i loro strumenti di lettura, la cui manutenzione diventerà sempre più onerosa con il tempo.

D'altronde è imperativo distruggere fisicamente i supporti quando si desidera eliminare totalmente un'informazione. Non bisogna accontentarsi, quando la modalità di incisione lo permette, di cancellare i dati per distruggere l'informazione.

V.3. La conservazione a lungo termine dei dati

V.3.1. Migrazione dei dati

Ci sono due tipi di migrazioni che conviene distinguere:

- la migrazione che permette di cambiare supporto fisico di stoccaggio senza toccare la sequenza dei bit che contengono l'informazione;
- la migrazione che conduce a cambiare il formato o la codifica dei dati.

Ogni applicazione è costruita in effetti su linguaggi, programmi e sistemi di gestione che sono oggetto di modifica e di arricchimento permanenti. La maggior parte di questi, tuttavia, altera le condizioni di restituzione dei dati e delle informazioni a tal punto, in alcuni casi, da renderli inaccessibili e/o inutilizzabili.

È per questo che la soluzione raccomandata da un quarto di secolo per risolvere questo problema è la migrazione dei dati.

Quest'operazione consiste nel separare dati e informazioni di ogni strumento software servito per la loro costruzione e nel convertirli in un linguaggio standardizzato, che offra garanzie di perennità, come l'ASCII. Ma a termine la migrazione dei file potrebbe farsi verso il nuovo codice che è UNICODE.

La tecnica della migrazione assicura la perpetuità dei dati e delle informazioni. Essa impone un'attenzione tecnologica permanente. Ciò implica sia che gli Archivi contino fra i loro effettivi persone che conoscono

l'evoluzione dei linguaggi e degli strumenti, sia che essi possano trovare tali competenze presso il servizio informatico del soggetto da cui dipendono o al quale sono legati.

I dati così separati dalla loro formattazione originale non hanno più, una volta portati su un altro programma, il loro aspetto e la loro presentazione originale.

La conservazione elettronica dei dati con la tecnica della migrazione porta dunque a un risultato molto differente da ciò che avviene con la carta e con gli altri supporti analoghi: per questi ultimi si provvede a conservare allo stesso tempo l'informazione e il modo in cui questa è stata, fin dall'origine, formattata. E questa formattazione spesso costituisce essa stessa una fonte d'informazione. Chiunque consulterà documenti o dati che hanno costituito oggetto di migrazione sarà dunque tributario degli strumenti di ricerca e soprattutto della descrizione dei metadati che conterranno quelli, per ricostituire l'aspetto iniziale dei documenti e la percezione che potevano averne produttore(i) e utenti.

Si può dunque dire che la migrazione è in una certa misura irreversibile come la separazione dei dati dal loro contesto tecnologico di creazione. Invece una migrazione da uno stato di file ASCII a uno stato di file UNICODE è in seguito reversibile, il che è un elemento importante di garanzia dell'integrità dei dati.

In altri casi, quando il modo di codificazione è specifico per un costruttore e per un sistema di utilizzazione e si desidera far migrare i dati verso un modo di codificazione standard, sarà impossibile assicurare la reversibilità dell'operazione.

Esempio dell'irreversibilità di migrazione di un modo di codificazione

Il CNES manipola dei numeri con una grandissima precisione. È stato costretto a procedere a migrazioni perché il modo di codificare i numeri reali utilizzato (per esempio 12398,9865490) era proprietario e era preferibile passare a una codificazione standardizzata. In ragione della misura dei numeri in macchina, non è stato possibile, durante quest'operazione, garantire che l'ultimo decimale fosse identico prima e dopo la migrazione.

V.3.2. Le microforme COM (Computer Output Microform)

E' possibile produrre microforme direttamente presso i dati binari provenienti da un computer: microfilm 16 o 35 mm e più comunemente microfiche.

Esistono due categorie di macchine di produzione: COM alfanumerica, limitata alla trascrizione dei segni dalla scrittura all'esclusione di schemi, schizzi e fotografie, e COM grafica, adatta a riprodurre e a registrare tutte le forme grafiche. È in seguito possibile ricostituire l'oggetto numerico a partire dalla microforma COM.

La tecnica COM assicura dunque una copia di sicurezza al documento digitale dal momento che ogni garanzia sull'identità e sull'integrità dei documenti è ben assicurata dallo strumento (nessuna possibilità di intervento umano nel trasferimento dalla forma digitale a quella analogica per esempio).

Invece disporre di microforme, la cui durata di vita in buone condizioni climatiche e materiali è di un centinaio di anni e che hanno così il vantaggio di essere leggibile senza intervento di computer, non dispensa dal prestare attenzione alla buona conservazione della forma digitale. La microforma priva delle facilità di trattamento e di ricerca che presenta il digitale. Essa ha necessità di apparecchi di lettura relativamente poco usati.

Tuttavia ciò può essere una risposta per tipologie ben mirate. Dati sottomessi a un'esigenza forte in materia di sicurezza e di fissazione dell'informazione e la cui distruzione farebbe correre dei rischi finanziari e giuridici meritano che si consideri il ricorso a questa tecnologia più onerosa. In questo caso bisognerà far attenzione alla perfetta conservazione delle microforme senza tuttavia dimenticare quelle dei dati digitali.

La SNCF per esempio utilizza la tecnica COM per la conservazione dei piani dei convogli TGV, prodotti e conservati ugualmente nella loro forma digitale.

V.4. I locali di archiviazione

Come un servizio d'archivio ha magazzini di stoccaggio per i supporti tradizionali, così deve aver previsto spazi di conservazione per i dati elettronici, un servizio d'archivio elettronico. È necessario prevedere la loro capacità per un dato lasso di tempo (fra 8 e 10 anni è ragionevole), in funzione della produzione suscettibile di essere archiviata e della politica e dei mezzi di conservazione definiti dall'istituto. Questi mezzi tecnici di conservazione concorrono ad assicurare l'integrità dei dati e a garantire il loro accesso attraverso il tempo.

È sufficiente oggi riservare nei locali adattati alla conservazione su supporto tradizionale un locale o uno spazio ben protetto, senza polvere, senza luce e con un campo magnetico il più debole possibile, evitando in particolare la prossimità di motori o di trasformatori elettrici, una temperatura di circa 20°,

un'umidità relativa del 40%. I supporti saranno condizionati in scatole, che evitino ogni deformazione o rigature dello strato di policarbonato.

VI. Mettere a disposizione i dati

È con la messa a disposizione dei dati che possono essere cambiati l'immagine e il ruolo dei servizi d'archivio nella società.

La disponibilità dei dati digitali costituisce un fattore determinante nell'accrescimento dell'attività di comunicazione. La comunicazione può ormai essere fatta a distanza e essere accessibile nello stesso tempo a un numero indeterminato di persone. Essa può anche essere materialmente assicurata da prestatori di servizio o da altre istituzioni diverse da quelle che assicureranno la conservazione di un altro esemplare fuori linea dei dati. Il fatto che i documenti saranno consultati da un gran numero di persone in ragione di questo miglioramento dell'accessibilità, in particolare quando i dati sono in linea, può spingere la tutela dei servizi d'archivio ad assicurare a questi delle risorse di bilancio e umane più importanti, risorse che potranno essere dedicate a un trattamento dei dati elettronici più pesante di quello dei documenti su carta.

I servizi, che sapranno accettare la sfida di una comunicazione immediata dei documenti legalmente comunicabili più richiesti, risponderanno veramente alle esigenze della società dell'informazione, per la quale elettronica è sinonimo di disponibilità.

I sistemi di comunicazione dei dati dovranno dunque tener conto automaticamente dell'accesso e della gestione dei diritti. I servizi d'archivio avranno interesse a mettere in linea i documenti più richiesti e ad assicurare dei servizi di manutenzione 24 ore su 24 dei relativi server.

L'amministrazione produttrice o il ricercatore ritroveranno traccia dei dati grazie agli strumenti di ricerca archivistica che terranno conto, eventualmente in modo automatico, della documentazione e dei metadati elaborati durante la produzione del documento.

L'amministrazione e il grande pubblico procederanno, come di consueto, alla domanda di comunicazione che potrà rilevare dal campo delle teleprocedure.

VI.1. Gli strumenti di ricerca

La descrizione è una delle possibilità più importanti del lavoro di preservazione a lungo termine dei dati elettronici. L'archivista deve essere cosciente che il lavoro di analisi e di elaborazione di strumenti di ricerca è molto più lungo e delicato per l'informazione digitale che per documenti su supporto cartaceo, che supportano maggior approssimazione. L'analisi di un'applicazione non può essere redatta in qualche riga.

Nel caso del documento elettronico in effetti non è l'unità documentaria di qualunque misura sia che si descrive, ma i dati che essa contiene. Nel caso di una base di dati, il contenuto di ognuno dei dati è spiegato nel dizionario, le sue relazioni con gli altri sono descritte. L'archivista farà attenzione anche a descrivere il contesto di creazione e di vita dell'informazione della sua raccolta, le sue lacune allo scopo di chiarire al lettore lo stato nel quale essa gli è fornita.

Sarebbe dunque preferibile che questa descrizione fosse in parte fatta molto a monte durante la produzione dei dati. La definizione di questi metadati risulta da una collaborazione tra archivista e produttore. I bisogni di ricerca dell'amministrazione o dell'impresa sono allora perfettamente presi in considerazione come quelli del grande pubblico.

Questa descrizione può essere fatta con gli strumenti usuali dell'istituto e anche su supporto tradizionale, ivi compresa la carta. Essa deve essere conforme alle raccomandazioni della norma ISAD (G), versione 2 (settembre 2000), pubblicata dal Consiglio internazionale degli archivi.

A termine, si potrà far attenzione che una parte dei metadati possa alimentare automaticamente lo strumento di ricerca. In effetti, nel caso di dati codificati o incapsulati nel formato XML, è del tutto possibile che uno di essi serva ad alimentare un sistema di descrizione anch'esso scritto in XML, per esempio ricorrendo alla DTD (EAD) (*Document Type Definition/Encoded Archival Description*).

Questa DTD compatibile con ISAD (G) è lo strumento elettronico attualmente più adatto alla descrizione gerarchica dei contenuti.

È possibile richiedere, in un quaderno delle spese di gestione elettronica di documenti, che certi elementi alimentino automaticamente un catalogo o degli strumenti di ricerca, in linea o no. Si può anche avere, grazie a un elemento codificato in un tag XML, un legame dello strumento di ricerca con il documento elettronico stesso (per esempio una base di dati o dati di una base di dati).

È da rimarcare che certe basi costituite dall'amministrazione possono essere considerate come strumenti di ricerca, perché esse danno accesso ad altri documenti, per esempio fascicoli. Questi strumenti di ricerca, contrariamente alla pratica archivistica, non sono sempre immediatamente comunicabili, perché essi possono contenere informazioni confidenziali, soprattutto nominative. Si baderà dunque che lo strumento di ricerca disponibile al grande pubblico sia ripulito da questi elementi non comunicabili, finché il termine della comunicabilità non sia scaduto.

Un caso specifico: il Servizio di trasferimento e di archiviazione dei file (STAF) del Centro Nazionale di Studi Spaziali (CNES)

L'archiviazione dei progetti comprende generalmente i dati stessi come anche i file descritti da questi dati. Lo STAF può gestire un legame semantico tra i dati e la loro descrizione, ma molto spesso sono gli utilizzatori stessi che preferiscono gestire in proprie basi questi legami.

L'amministratore del sistema che ha una panoramica orizzontale su tutti i progetti crea una radice, poi l'amministratore del progetto denomina il progetto, crea profili di utenti. Così ogni progetto ha la sua struttura (radice, repertorio, file) e alimenta il catalogo.

Non si tratta di una descrizione nel senso in cui gli archivisti l'intendono abitualmente. Invece questi elementi fanno proprio parte dei metadati di descrizione e sono automaticamente integrati con lo strumento di ricerca.

VI.2. Comunicazione dei dati

Secondo i tipi di dati, una mediazione del servizio di archiviazione sarà necessaria per trasmetterli, nel caso in cui questi, richiesti di rado, saranno stati conservati soltanto fuori linea.

Il pubblico, secondo le scadenze di comunicabilità, potrà o no accedere direttamente ai dati; dopo aver trovato una prima pista di ricerca nella descrizione dei dati, sarà allora informato sul loro contenuto più dettagliato dal servizio di comunicazione; potrà così se necessario cominciare una richiesta di deroga nel caso di dati gestiti dai servizi pubblici secondo le leggi in vigore.

Ma se in tutti i casi è utile avere una conservazione dei dati grezzi sui supporti più pertinenti, si avrà anche interesse a farli montare dal servizio informatico competente più vicino su qualunque programma di gestione di basi utilizzato correntemente dall'istituto di conservazione a lungo termine e a mantenerli su server, dal momento in cui i dati sono suscettibili di essere richiesti spesso,

allo scopo di fornire l'informazione in tempo reale al pubblico e all'istituto produttore. Questo è particolarmente vero per le relazioni fra il servizio d'archiviazione e i servizi produttori, che non verseranno i loro documenti elettronici negli archivi pubblici se questi non renderanno questo servizio istantaneo durante la ricerca amministrativa.

Questo significa che i servizi di conservazione a lungo termine potranno portare i dati in un nuovo sistema di gestione di basi e renderli interrogabili non secondo lo stesso modo iniziale dell'istituto produttore dei dati, ma secondo le modalità giudicate più utili ai ricercatori di qualunque tipo essi siano (amministrazione produttrice e altro, ricercatore).

Bisognerà anche avere un servizio di rilascio al ricercatore di dati a sua scelta grezzi su supporti standard di mercato; il ricercatore potrà in seguito installarli o farli installare sui sistemi di gestione a sua scelta.

Se si è indotti a considerare una tariffazione della fornitura di dati elettronici alle altre persone piuttosto che al servizio produttore, ci si conformerà al decreto n° 2001-493 del 6 giugno 2001¹⁷ preso per l'applicazione dell'articolo 4 della legge n° 78-753: "In occasione del rilascio del documento, spese corrispondenti al costo di riproduzione e, caso eventuale, di invio di questo e che costituiscono una remunerazione per servizi resi possono essere messi a carico del richiedente. Per il calcolo di queste spese si tiene conto, con l'esclusione dei carichi di personale risultante dal tempo dedicato alla ricerca, alla riproduzione e all'invio del documento, del costo del supporto fornito al richiedente, del costo di ammortamento e di funzionamento del materiale utilizzato per la riproduzione del documento così come il costo dell'affrancatura secondo le modalità di invio postale scelte dal richiedente".

Gli stessi dati elettronici possono essere accessibili nello stesso istituto in diversi modi o in molte istituzioni in diversi modi. La descrizione dei dati deve tener conto di queste differenti soluzioni e delle molteplici localizzazioni che essa implica. Dei dati possono essere messi in linea sia nello stesso sistema e nella stessa base di quando essi erano presso il loro produttore, sia messi grezzi su supporti di archiviazione fuori linea che possono essere essi stessi di molti tipi (un'istituzione conserverà a volte gli stessi dati su un disco e su una cassetta, perfino su un supporto analogico come microforme COM). Si deve tener conto nella descrizione del fatto che questi diversi supporti possono essere sotto una stessa responsabilità intellettuale e scientifica e non essere conservati nello stesso luogo fisico (così un servizio di archivi dipartimentali potrà conservare nei suoi locali i dati in linea su disco mentre gli stessi dati saranno conservati su cassette in un armadio assicurato dal

¹⁷ Il decreto si riferisce alle modalità di comunicazione dei documenti amministrativi.

servizio informatico del Consiglio generale o da una società di sottotrattamento ...).

VII. Studi di casi

Si parlerà in questo capitolo di alcuni sviluppi concreti dei principi generali enunciati nel capitolo precedente per documenti e applicazioni elettronici provenienti dai servizi d'archivio e le cui procedure e tecniche di conservazione a lungo termine sono già state testate e collaudate.

VII.1. Basi di dati

Una base di dati è costituita dal raggruppamento di molti file e di tavole in una sola applicazione e può dunque essere relativamente complessa.

Esistono molte tecnologie per la gestione di basi di dati. Il linguaggio oggi più diffuso, SQL, permette di manipolare molto facilmente le basi di dati concepite con un SGBD-R (sistema di gestione di base di dati relazionale) (aggiunta, soppressione, selezione di dati) e di fare delle ricerche.

Certe basi sono messe in opera per un bisogno puntuale: è il caso dei file di inchieste e di statistiche. Una tale base forma un tutto semplice da identificare, le cui procedure di versamento possono essere molto facilitate al di là di una scadenza di utilità amministrativa facile da determinare con il produttore. Invece la maggior parte delle basi di dati, per esempio quella di gestione del personale, sono vive, vi sono aggiunti quotidianamente dati o ne sono estratti e cancellati. Il compito dell'archivista e del produttore sarà dunque più delicato, specialmente per trovare i punti periodici di validazione e di fissaggio dei dati.

Come procedere? Il servizio produttore sa generalmente per quanto tempo i dati registrati gli saranno utili. Tutto il problema proviene in generale dal fatto che è stato previsto nell'istituto produttore di sopprimere i dati alla fine di questa scadenza, quando i server non hanno la capacità di stoccare altra informazione e che l'archivista non è al corrente di questa procedura. È dunque indispensabile che egli solleciti il versamento regolare dei dati che avrà selezionato con il servizio produttore come degni di essere conservati a lungo termine.

L'oggetto d'informazione deve obbligatoriamente avere al momento dell'archiviazione una struttura determinata e validata.

Quale può essere la fase determinata e validata di una base di dati viva?

Ci sono molte risposte a ciò, secondo il contenuto di questa base e il tipo dei dati:

- l'archiviazione di una fotografia della base può essere fatta con una periodicità determinata;
- la prima archiviazione prende la forma di una fotografia della base, mentre le ulteriori archiviazioni raccolgono unicamente i dati nuovi dopo la data di archiviazione precedente;
- si archivieranno nei due casi anche i dati che sono ritirati dalla base; l'estrazione dalla base viva è un criterio di archiviazione in sé;
- il produttore può determinare altri tipi di fasi di validazione, di recinzione di parti della base (per esempio fascicoli all'interno di essa) in relazione con l'archivista e prevedere l'archiviazione secondo questi criteri di validazione.

Se si decide di conservare una fotografia periodica della base effettuando per esempio un versamento annuale della sua totalità, ogni versamento riprenderà l'integralità dei dati del versamento precedente e vi aggiungerà soltanto le modifiche o i dati nuovi intervenuti nell'anno in questione. Si archivierà dunque in questo caso una grande massa di dati di cui una piccola parte soltanto rappresenterebbe delle informazioni nuove veramente da conservare. Il vantaggio di questo metodo è che si è sicuri di archiviare, senza alcuna perdita, ma con ridondanze, tutti i dati della base.

Invece questo metodo non permette di risolvere due difficoltà maggiori:

- le conseguenze delle modifiche con effetto retroattivo che determinano conseguenze sul periodo cronologico precedentemente archiviato nella situazione amministrativa di un qualunque agente (nel caso di una base di gestione di personale), a seguito di modifiche regolamentari. La pertinenza e la validità dei dati archiviati in ogni versamento non possono dunque essere assicurate. Il ricercatore che si trova di fronte ai diversi versamenti non potrebbe fare comparazioni valide e non sarebbe in condizione di verificare la validità di informazioni contraddittorie secondo i versamenti (per esempio per uno stesso agente in una stessa data);
- i bisogni del servizio versante. Poiché ogni versamento rappresenta effettivamente l'integralità delle informazioni contenute nella base alla data nella quale è effettuato, sarebbe essenziale che il servizio versante non ne distruggesse alcuna da un anno all'altro. Dunque, se il servizio versante prova l'esigenza di archiviare i dati, è per il desiderio che siano conservate le informazioni da cui è obbligato a non separarsi per mancanza di posto fisico.

È dunque indispensabile nel caso di una base i cui dati vivono potenzialmente in permanenza, definire altri criteri di estrazione: per esempio per una base di gestione di personale, le informazioni relative agli agenti che hanno lasciato definitivamente l'istituto durante la periodicità decisa per l'archiviazione (per esempio durante l'anno civile passato, se l'archiviazione ha luogo tutti i 2 gennaio). Si può anche decidere, per incrementare la validità dei dati, di non procedere a quest'archiviazione se non alla fine di tre anni, per esempio di archiviare nel 2010 i dati concernenti il personale che ha lasciato l'istituto nel 2007, allo scopo di permettere le modifiche che possono intervenire per regolarizzare situazioni dopo la partenza dell'agente. Una volta versati questi dati al servizio di archiviazione definitivo, il servizio che versa può eliminarli totalmente dalla sua base. Questa procedura è molto più soddisfacente. Bisogna provare a mantenere nella gestione a lungo termine del documento elettronico le fasi transitorie che si utilizzano per la carta. Procedere a un versamento prima della scadenza di queste fasi-tampone che permettono alla riflessione di maturare potrebbe far perdere in fin dei conti del tempo.

Si procederà nello stesso modo per esempio per fascicoli di giustizia e si potrebbe operare in questo caso l'archiviazione definitiva per fascicoli esauriti tre anni dopo la loro chiusura.

Quando si archivia una base mettendo grezzi i dati tramite la loro migrazione, li si conserva da soli senza i programmi che sono serviti per crearli e per utilizzarli. Non si conserva mai la base come si è presentata ai suoi utilizzatori iniziali. È dunque necessario tracciare di nuovo negli strumenti di ricerca che la descrivono la sua struttura d'origine che sparisce con i programmi di gestione dei dati, i tipi di diritti che permettevano di accedervi. Questa raccolta di metadati indispensabili per comprendere come era organizzata in origine la base non permetterà tuttavia di ricostruirla a colpo sicuro nel suo stato iniziale.

I versamenti di basi di dati contengono molte parti distinte:

- i prospetti contenenti i dati in linguaggio di codificazione standardizzato, per esempio l'ASCII;
- la struttura della base che comprende la struttura dei differenti prospetti (liste e tipi dei campi) e le relazioni che li legano;
- i codici e la loro stesura in chiaro per ogni tipo di dati codificato.

Ogni prospetto costituisce l'oggetto di un file durante il versamento. Ma se si desidera rimontare rapidamente la base in un nuovo sistema di gestione per esempio per la comunicazione, si potrà anche farsi consegnare un file contenente in seguito i dati in ASCII delimitato.

Tutti i dati sono peraltro descritti e esplicitati in un dizionario dei dati. Esso costituisce come i prospetti o dizionario dei codici e la struttura della base, una parte della documentazione o metadati che sono recuperati dai servizi di archiviazione a lungo termine sotto forma elettronica.

Per i prospetti dei codici, si starà attenti al fatto che uno stesso codice può cambiare significato con il tempo per molte ragioni. L'ideale sarebbe che una tale situazione non si verificasse e che i gestori di basi facessero attenzione a non far verificare questo caso. Ma è inevitabile che l'archivista vi si confronti un giorno o l'altro. Per esempio, nel caso di base di gestione di personale, sarà in ragione della creazione di corpi di nuovi funzionari che si sostituiscono ai vecchi, della trasformazione di status o di denominazione. Importa conservare tutti i valori presi da uno stesso codice. Si tratterà dunque di stabilire la lista dei codici di una base di dati con l'integralità dei significati differenti che afferiscono a ciascun codice, precisando la loro data d'inizio e di fine di validità e questo per tutti i prospetti necessari alla comprensione delle basi di dati. Perciò sarà necessario trovare all'interno dei servizi produttori le persone che conoscano i codici vecchi e obsoleti presenti nei file in corso di archiviazione.

Si può evitarlo spingendo i servizi informatici che gestiscono questi prospetti a conservare la loro cronistoria, cioè la traccia delle modifiche, soppressioni, aggiunte e loro date. Le nuove versioni dei prospetti che accompagnano i nuovi versamenti saranno allora trasmesse agli Archivi, senza che questi debbano ricostruire la cronistoria di questi prospetti. Questo lavoro di ricostituzione è in effetti molto pesante e rischia di non essere fatto o di essere fatto a posteriori.

Si è visto che si poteva per le basi di dati concepire due modi paralleli di conservazione a lungo termine nei servizi di archiviazione storica, in funzione dei bisogni di comunicazione, modi che permetteranno da una parte di conservare i dati grezzi, dall'altra li manterranno accessibili rapidamente per la ricerca storica o amministrativa.

È così che si è proceduto con la Direzione degli Archivi di Francia all'interno dei servizi dipartimentali della corona parigina per la Nouvelle Chaîne Penale di Parigi e della regione parigina.

Invece non sembra ragionevole voler conservare basi di dati in formato cartaceo. Non c'è altra soluzione per conservarle che farlo in modo elettronico. L'investimento per concepirle e per farle vivere durante un lasso di tempo più o meno lungo sarebbe totalmente perduto se si praticasse così. Il valore della base di dati è la grande capacità di ricerca e di comunicazione possibile grazie ad essa.

Esempio: l'archiviazione della Nouvelle Chaîne Pénale di Parigi e della regione parigina (NCP)

Nel quadro di sperimentazioni sull'archiviazione elettronica, la Direzione degli Archivi di Francia ha deciso alla fine del 2000, su suggerimento del servizio degli archivi del Ministero della giustizia, di procedere a un test di archiviazione su un'applicazione del settore penale. Quest'applicazione chiamata "nouvelle chaîne pénale" è utilizzata nei tribunali di prima istanza di Parigi, Nanterre, Bobigny, Créteil, Pontoise, Evry e Versailles.

Il test verteva da una parte sulle procedure da mettere in opera per l'archiviazione di una base viva, mentre gli Archivi Nazionali avevano esperienza solo di basi di indagini statistiche puntuali, rigide, dall'altra parte sulle implicazioni che una tale operazione poteva avere per un servizio d'archivio territoriale.

Essendo questa base uno strumento d'accesso ai fascicoli giudiziari stessi, sarebbe stato deplorabile praticare soltanto un'archiviazione dei dati grezzi, in ASCII. Gli Archivi di Francia hanno dunque realizzato, su un programma di gestione di base di dati per il grande pubblico, un'interfaccia d'interrogazione che permette di fare rapidamente delle ricerche, senza dover rimontare i dati su un software, come avviene con le modalità di archiviazione tradizionale dei documenti digitali.

I metadati della "Chaîne Pénale" di Parigi e della regione parigina

Ecco basandosi sui tipi di metadati recensiti nel quadro del sottocapitolo IV.2, costituire i metadati, il che è per una base di dati come quella della Chaîne Pénale di Parigi e della regione parigina, a fianco agli Archivi di Parigi: 2157 W.

Metadati di conservazione:

- *nome dei file: affaires.dat; personnes.dat; textes.dat ;*
- *tavole di riferimento:*
 - *coraff.txt (concordanza tra antica e nuova nomenclatura delle nature di affari);*
 - *elst.txt: elementi di struttura;*
 - *elstloc.txt: elementi di struttura locali;*
 - *evt.txt: tipo degli avvenimenti che avviano l'azione pubblica;*
 - *motevt.txt: motivo degli avvenimenti;*
 - *nataff.txt: natura d'affari;*
 - *natine.txt: natura delle infrazioni;*
 - *rôle.txt: ruolo delle parti;*
 - *service.txt: nomi dei servizi che hanno trattato gli affari;*
 - *stajuri.txt: stato giuridico delle persone morali.*
- *Le tavole dei codici che datano dal 2001 sono state consegnate a parte.*
- *Sistema di utilizzazione e SGBD-R: i programmi su Geos 7; basi di dato Oracle versione 7 su server UNIX.*

Metadati di descrizione:

- *intitolazione del documento: “Registrazione sulla Nouvelle Chaîne Pénale (NCP) degli affari penali del Tribunale di prima istanza di Parigi: affari classificati senza seguito e affari giudicati ripresi dall’antica applicazione Ufficio di ordine penale (BOP), 1973-1986”;*
- *contesto: “Applicazione sviluppata presso il Ministero della giustizia, nel settore penale, chiamato “nouvelle chaîne pénale” e utilizzata nei tribunali di prima istanza di Parigi, Nanterre, Bobigny, Créteil, Pontoise, Evry e Versailles”;*
- *servizio produttore: Ministero della giustizia, Tribunale di prima istanza di Parigi, Procura;*
- *data di creazione della NCP: 1993 (primo sito impiantato a Pontoise)*
- *formato dei dati: ASCII delimitato;*
- *software di cattura: programma sotto UNIX di estrazione dei dati dalla base Oracle, richiesta SQL, trattamento per lotti.*

Metadati di gestione:

- *data di versamento: 6 dicembre 2001;*
- *servizio versante: Ministero della giustizia, Direzione del l’amministrazione generale e delle attrezzature, Centre de prestations régionales di Grigny (dove sono immagazzinati i dati del modulo di archiviazione della NCP);*
- *indirizzo di localizzazione: Archivi di Parigi;*
- *supporti: dati a terra e tavole di riferimento su cassetta DAT 4 mm e copia su CD-Rom; dati delle tavole rimontate su Access consultabili dal personale su un posto dedicato agli Archivi di Parigi e copia su un CR-Rom;*
- *volumetria: 300 megabyte;*
- *comunicabilità: 100 anni;*
- *conservazione: illimitata.*

VII.2. I documenti provenienti dalla burocratica e la posta elettronica

VII.2.1. I documenti provenienti dalla burocratica

Quasi tutti i documenti d’ufficio sono creati da qualche anno grazie a uno strumento elettronico; di nuovo c’è che essi non sono più sistematicamente stampati su carta. L’uscita su supporto cartaceo risponde ancora a due bisogni: un bisogno ergonomico, di semplice comodità, che fa sì che certi preferiscano lavorare e soprattutto correggere su uscite cartacee, e che si apprende meglio se si continua ad avere una lettura lineare; un bisogno giuridico o di prova che fa portare un segno evidente di convalida, come una segnatura manoscritta, sui documenti di carta prima che siano diffusi e che spinge ad archivarli in questa forma per conservare la traccia di questa convalida.

Una parte della produzione d’ufficio è informale e ufficiosa. Può consistere in un’informazione scambiata unicamente sulle reti da un piccolo numero di persone spesso senza alcun segno di convalida. Essa può consistere anche in brutte copie, annotazioni, versioni preparatorie non convalidate il cui valore

storico può essere forte perché esse illuminano sulla gestazione dei documenti ufficiali e convalidati.

Questi file restano immagazzinati un certo tempo sui dischi duri dei loro redattori e sono, in generale, al momento della partenza di quest'ultimo, modesti. La confusione attuale è tanto più grave quanto più si sviluppa il lavoro condiviso, in rete. L'istruzione di un fascicolo è oggi spesso ripartita tra più persone che possono appartenere a servizi differenti. Sia i *software* non gestiscono facilmente e automaticamente la sopravvivenza delle differenti versioni dei documenti, sia gli utilizzatori non sfruttano tutte le potenzialità dei loro strumenti informatici.

Si assiste da qualche anno alla generalizzazione degli strumenti di burocratica, ma gli agenti amministrativi non hanno veramente coscienza che questi li hanno fatti entrare in un universo nuovo, quello della gestione elettronica dei documenti. I responsabili fanno di conseguenza raramente una riflessione globale sulla produzione dei loro servizi, mentre questa è oggi molto complessa, essendo la carta ancora utilizzata insieme con il supporto elettronico.

L'archivista e tutti gli altri responsabili della gestione documentaria devono sensibilizzare gli agenti al fatto che l'informazione che essi producono deve essere trattata nella sua globalità, che essa deve essere classificata in fascicoli di carta o in file elettronici e per uno stesso affare perfino nelle due forme.

Si tratterà così di determinare in funzione del valore probante dei documenti da una parte e del loro interesse storico dall'altra, quali sono le versioni del documento da conservare, proveniente da quale server, da quale posto di lavoro, da quale fascicolo cartaceo. I metodi di lavoro restano invariati, ma l'elettronica introduce uno strato di complessità nella produzione.

Un documento cartaceo isolato, non inserito in un fascicolo, corre il rischio di essere definitivamente perduto. Un fascicolo che non è integrato in un piano di classificazione è più difficile da ritrovare rispetto a quello che lo è. Questi due principi di base sono applicabili ai file burocratici.

L'archivista o l'amministratore dei dati dovranno spiegare agli utilizzatori che il piano di classificazione utilizzato per i documenti cartacei vale per i repertori elettronici. Esso permetterà di strutturare la produzione dei file digitali, il che si rivela indispensabile per la loro ricerca, perché dà loro un senso, e per la loro archiviazione.

Ogni file burocratico deve essere documentato dal suo autore per essere ritrovato facilmente tra gli altri con l'attribuzione di una parola-chiave (cioè

un'indicizzazione del documento, a partire dai termini del piano di classificazione e da una lista d'autorità), con l'indicazione del nome dell'autore, del destinatario, del titolo, dell'oggetto. Se quest'identificazione non è fatta per il documento, essa dovrà esserlo imperativamente per il repertorio nel quale il documento è classificato.

Raccomandazioni per la tenuta della documentazione del progetto GEREHMI (strumento di gestione delle risorse umane) del Ministero della giustizia: estratto dal piano di sicurezza qualità del progetto

“Il responsabile funzionale è responsabile della buona tenuta della documentazione. Può appoggiarsi per questo compito sul segretariato del dominio più particolarmente incaricato della tenuta della documentazione cartacea.

La documentazione ha un'importanza fondamentale: è infatti lo strumento di comunicazione e di dialogo tra i membri del gruppo e i soggetti esterni che intervengono (membri dei comitati, utenti, ...). Essa permette di garantire la perennità delle informazioni all'interno del progetto.

Un meccanismo normalizzato di identificazione dei documenti è stato approntato in modo da assicurare l'efficacia della gestione della documentazione. I documenti sono dunque referenziati.

Ogni documento ha un ciclo di vita, cioè attraversa un certo numero di stati.

	Stato del documento	Descrizione dello stato
<i>T</i>	<i>Lavoro</i>	<i>Il documento è in corso di elaborazione da parte dell'autore.</i>
<i>P</i>	<i>Provvisorio</i>	<i>Il documento è terminato, in attesa di essere pre-convalidato.</i>
<i>C</i>	<i>Pre-convalidato</i>	<i>Il documento è approvato dagli altri membri del gruppo, soggetti esterni che intervengono ... La pre-convalida del documento può essere opzionale.</i>
<i>V</i>	<i>Convalidato</i>	<i>Il documento è approvato dalle persone abilitate e acquista valore di riferimento all'interno del progetto. Diviene applicabile.</i>
<i>A</i>	<i>Archiviato</i>	<i>Il documento non è più consultato regolarmente, ma resta una traccia della sua esistenza.</i>

La gestione dello stato di un documento si fa tramite un quadro di cui si dà un esempio qui di seguito:

Nome del documento	Oggetto	Versione	Data della versione	Redattore	Stato	Archiviato su carta
<i>PAQ</i>		<i>V1.1</i>	<i>17/2/2000</i>	<i>JF LEBEC</i>	<i>Convalidato</i>	<i>X</i>

GEREHMI						
PAQ GEREHMI		V1.2	23/3/2000	JF LEBEC	Lavoro	

Questo quadro sarà aggiornato regolarmente (cioè in occasione di ogni cambiamento di stato di un documento ufficiale progetto) dal redattore (responsabile) del relativo documento.

Creazione di un nuovo file. I seguenti punti dovranno essere rispettati:

- *sempre lavorare sul server (e non su una postazione di lavoro) in modo che il documento sia sempre disponibile per l'insieme del gruppo e che costituisca l'oggetto di tutela quotidiana). Prevedere una V0 [versione zero) per i documenti di lavoro;*
- *rispettare l'organizzazione del server: trovare il repertorio pertinente. Se non esiste, valutare con il responsabile della qualità la creazione di un nuovo repertorio;*
- *dare un nome esplicito al nuovo file (nessuna limitazione sulla lunghezza) e conforme alle norme preesistenti. Segnalare obbligatoriamente le proprietà e rispettare i suffissi stabiliti di default (.doc, .xls, .wdp, .ppt);*
- *informare il segretariato del progetto della creazione di un nuovo fascicolo per l'aggiornamento della lista dei documenti.*

Modifica di un file esistente

Aggiornare eventualmente lo stato di un fascicolo caratterizzato dal suo numero di "versione-revisione":

- *il numero di versione cambia se modifiche di fondo condizionano il fascicolo;*
- *il numero di revisione cambia se modifiche di forma (o adattamenti minori) condizionano il fascicolo.*

Mettere a livello (eventualmente):

- *lo stato del fascicolo rilevato sulla pagina di guardia;*
- *il fondo pagina;*
- *le proprietà".*

Infine, dato che formati diversi possono essere immagazzinati in uno stesso repertorio di burotica, l'archivista raccomanderà di utilizzare, per la creazione o per la conversione dei file, i formati più duraturi. Una difficoltà sta nel fatto che gli strumenti burotici attuali non permettono una facile conversione in XML. In attesa di questo, il formato PDF è un candidato relativamente buono per l'archiviazione dei documenti provenienti dalla burotica.

L'archivista può tentare, se ottiene i mezzi per mettere a posto un applicativo per l'archiviazione di questi file, di automatizzare i loro versamenti. Le scadenze di utilità amministrativa e la sorte finale dei file possono essere integrate nei loro metadati. Alla scadenza, possono essere prodotti dei file XML, che contengono i dati e i metadati che possono essere trasferiti con FTP al server di archiviazione dell'istituto di conservazione. La messa a punto di un tale dispositivo richiede la creazione di DTD specifiche per ogni tipo di file da archiviare. Questo lavoro può essere pesante nella misura in cui questi file sono molto spesso informali e non strutturati.

Esempio: l'archiviazione dei documenti burocratici nel Ministero della giustizia

Un lavoro di sensibilizzazione è stato condotto presso il Ministero della giustizia con l'Ispezione generale dei servizi giudiziari nella misura in cui gli ispettori incaricati di una missione precisa lavorano sistematicamente in rete e producono un numero impressionante di documenti elettronici, con diversi formati (tenuto conto di visite, colloqui, note ...) preparatori al rapporto finale, che non saranno mai pubblicati. Questi documenti quando erano prodotti su carta erano integrati nei fascicoli ed erano versati, mentre gli stessi documenti elettronici sono distrutti, una volta compiuta la missione o che l'ispettore è partito. I file sono creati senza strutturazione iniziale, immagazzinati alla rinfusa in repertori che portano nomi poco chiari. Essendo stato richiesto all'Ispezione dal servizio d'archivio di elaborare un piano di classificazione per l'insieme dei documenti prodotti e ricevuti da essa, un piano di classificazione è stato prodotto ed è stato richiesto all'Ispezione di utilizzare questo piano non soltanto per procedere alla registrazione informatizzata di questi documenti, ma anche di servirsi delle rubriche del piano per strutturare gli uffici dei differenti ispettori sui loro posti di lavoro.

Il servizio d'archivio ha anche elaborato, con il servizio, un quadro di gestione "appoggiato" a questo piano di classificazione e ha raccomandato che, compiute le missioni, sarebbero archiviati tutti i documenti preparatori qualunque sia il loro supporto: i messaggi elettronici e le note e lavori redatti che non provengono dalla carta dovranno essere incisi su CD-Rom, essendo stati preventivamente convertiti i file nel formato PDF.

- 1) *Per l'arretrato, conviene indicare:*
 - *il nome del repertorio nel quale sono raccolti i diversi file relativi alla missione compiuta;*
 - *il nome e il numero dei file;*
 - *un titolo generale per il repertorio.*

- 2) *Per l'avvenire, bisognerà fornire per ogni file che compone il repertorio:*
 - *l'indicazione del codice del piano di classificazione utilizzato;*
 - *le parole chiave (tratte dalla lista d'autorità);*
 - *la data e un'analisi sommaria.*

Posto questo, la realizzazione sarà più complessa di quanto non sembrava: pesantezza d'integrare il piano di classificazione sulla rete dell'Ispezione come sui differenti posti di lavoro degli ispettori, trovandosi certi file su dischi duri; pesantezza della conversione e modalità della ripartizione dei compiti fra i servizi e il servizio d'archivio e mentalità da far evolvere, quanto alla presa di coscienza che i file elettronici di lavoro sono ugualmente archivi pubblici da versare.

VII.2.2. La posta elettronica

La prima tappa per conservare correttamente a lungo termine la posta elettronica è stabilire procedure chiare e logiche di registrazione come si pratica per la posta di carta con un oggetto, una data, un autore.

Il redattore dovrà anche indicizzare, anche sommariamente, ogni invio, poi strutturare nei repertori la sua posta. Per l'indicizzazione, sceglierà termini di preferenza predeterminati e controllati dall'archivista, il documentarista o il responsabile del Record Management dell'istituto produttore. Questa indicizzazione permetterà l'integrazione di ogni messaggio elettronico in un piano di classificazione. È dunque necessario avere un'architettura di conservazione parallela a quella della messaggeria per classificare la posta e i documenti associati e per procedere alla loro archiviazione.

I servizi potranno anche mettere a punto una politica comune di cassette delle lettere funzionali per grandi funzioni (se ne possono avere molte per servizio) per le quali il *software* di messaggeria prevedrà una funzione di archiviazione automatica.

La quasi totalità di questi metadati di descrizione e di registrazione può essere catturata automaticamente a partire dalla struttura stessa della posta elettronica. Bisogna che a monte l'archivista chieda che questa funzione sia ben presa in considerazione. L'incapsulamento, in un documento XML, per esempio, assicurerà loro allora una certa perennità.

L'esperienza della National Archives and Record Administration negli USA (NARA)

Sviluppando raccomandazioni sulle tecnologie necessarie per preservare i milioni di documenti di posta elettronica e dar accesso a utenti potenziali, la NARA ha chiesto al San Diego Super Computer Center – SDSC ... di trovare delle soluzioni tecniche compatibili con i principi archivistici del rispetto dei fondi e del rispetto dell'ordine originario dei documenti.

In risposta a queste esigenze il SDSC ha sviluppato un'architettura di gestione dell'informazione chiamata "Collection – based Persistent Object Preservation". Questa architettura incorpora un metodo che permette di trasformare i record in un formato indipendente dai software e dai materiali. Essa prende per base il modello OAIS.

L'esperienza condotta su molte collezioni della NARA, fra cui la posta elettronica, ha permesso:

- *di scegliere le collezioni di documenti;*
- *di trasformarli in un formato capace di durata;*
- *di metterli in un deposito di dati;*
- *di ricostruirli come collezione;*
- *di fare delle ricerche utilizzando una tecnologia differente da quella che era stata utilizzata per creare i documenti e immagazzinarli.*

Tutto ciò è stato realizzato in due giorni. Il successo di questa operazione ha spinto la NARA a proseguire le ricerche.

Molti problemi indipendenti dalle procedure e dalle tecniche di trattamento si pongono tuttavia per la conservazione a lungo termine della posta elettronica. Una parte importante di questa è in effetti di ordine secondario nei processi amministrativi (informazione, copia, ricezione) ed è anche nel novero di casi di ordine personale. Un buon numero di messaggi elettronici non sono firmati, non lo sono in modo da garantire l'identità del firmatario.

È per questo che l'archiviazione della posta elettronica pone anche le questioni:

- del carattere privato di questa posta. In quale misura l'archivista o il responsabile del *Record management* può intervenire nelle messaggerie per conservarle? In Francia molte decisioni di giustizia hanno condannato degli impiegati che accedevano alla messaggeria del loro impiego anche nell'ambito del lavoro. È il carattere privato della messaggeria che allora è stato preso in considerazione dal giudice;
- del valore probante di questi documenti. Gli inglesi hanno tentato di dare una risposta creando con la British Standards Institution un codice di pratica per l'ammissibilità legale di informazione immagazzinata in Electronic Document Management Systems – DISC PD008. Secondo loro, mettere in opera buone pratiche dettagliate in questi manuali, controllate da revisioni e rassicurate, potrebbe eliminare i dubbi dei tribunali. Nella pratica francese, questi codici di buone pratiche non sembrano sufficienti e il dibattito resta aperto anche se la norma NF Z 42-013 è una risposta diretta ma di aspetto più tecnico al DISC-PD008;
- della selezione e dello smistamento dei messaggi. Il produttore della posta dovrebbe poter precisare se la posta può essere archiviata. Sembrerebbe interessante avere processi di selezione volontaria avendo

in un'amministrazione che emette posta delle regole d'uso che dovranno essere rispettate da tutti gli agenti.

Quest'argomento particolarmente importante è dunque delicato per ragioni non tecniche. La Direzione degli Archivi di Francia approfondirà con le relative istanze la sua riflessione in questo campo nei tempi a venire.

VII.3. Risorse del web

Esistono molti manuali realizzati da gruppi di lavoro, manuali che sono tutti in inglese. La conservazione delle risorse web è, senza eccezione, assicurata oggi nel mondo da biblioteche che la considerano nella stessa logica di quella del deposito legale.

Tuttavia un numero sempre maggiore di documenti primari sono oggi pubblicati su internet. La raccolta di questi documenti fa parte dei compiti della Direzione degli Archivi di Francia e dei servizi che sono sotto la sua tutela allo stesso titolo di quella dei documenti amministrativi su altro supporto. I siti privati invece sono al di fuori del suo campo d'azione.

I servizi d'archivio potranno anche lavorare all'archiviazione dei siti intranet e extranet del loro istituto o degli istituti di cui essi gestiscono i documenti, siti che non possono essere toccati dai collettori automatici del web.

Due soluzioni sono prevedibili per l'archiviazione del web: sia archiviare ogni elemento componente di uno o più siti web, sia archiviare il (i) sito (siti) stesso in una fotografia o in uno strato in un dato momento (*snapshot*¹⁸ in inglese) e con una periodicità da definire.

La seconda soluzione è più semplice ma riguarda soltanto le pagine statiche accessibili a tutti, il che rappresenterebbe ancora la quasi totalità delle risorse web, in HTML, per il testo, o JPEG, per le immagini. Tutti gli altri elementi, dinamici o generati dinamicamente dal visitatore e che compongono sempre più i siti non possono invece essere archiviati in questo modo rudimentale. I programmi di raccolta automatica di siti web non riguardano gli strati profondi del web e non archiviano i documenti immagazzinati nelle basi di dati.

Non è possibile attualmente assicurare la conservazione a lungo termine dei documenti dinamici e di quelli che sono inseriti o dipendenti da programmi ausiliari, come Java Script.

¹⁸ Istantanea.

Per le basi di dati, per i formulari in linea e per le risposte che sono fornite, bisogna pensare a un'archiviazione specifica e programmata con i produttori delle relative pagine. Questo è più facile quando si decide di procedere all'archiviazione di un sito definito, preciso, di cui si controlla bene tutto il contenuto e per il quale si beneficia di un accesso alla gestione dei dati.

Si potrà anche decidere di archiviare periodicamente il flusso dell'utilizzazione in un lasso di tempo dato (per esempio una giornata). Ci si avvicinerà così a metodi di conservazione dell'audiovisivo praticati dall'INA¹⁹ che archivia periodicamente una giornata completa di diffusione.

Se si conservano a lungo termine soltanto i dati stessi, si farà attenzione a elaborare e a conservare una documentazione sufficientemente ricca, affinché il ricercatore sappia come questi dati sono stati presentati all'internauta e utilizzati da lui.

È dunque indispensabile mettere a punto una procedura di raccolta per servizio produttore di sito o di parte di sito web. Quando il sito web è gestito da molte persone e eventualmente su server diversi (come all'INRIA²⁰), l'operazione di raccolta centralizzata può rivelarsi più difficile. Si avrà allora interesse alla valutazione e all'archiviazione di ogni porzione di sito corrispondente a un dato produttore e a un dato server, tentando di mantenere, grazie a un piano di classificazione e a una descrizione archivistica strutturata, l'identità del tutto.

Un'operazione pionieristica è stata condotta dal 2001 dal Ministero del lavoro e della solidarietà.

Esperienza di archiviazione dei siti internet del Ministero del lavoro e della solidarietà

Un gruppo di lavoro, creato per iniziativa del centro di risorse del ministero poi pilotato dalla missione degli Archivi nazionali, si è insediato nell'estate del 2001. Questo gruppo è costituito per il ministero da informatici, da documentaristi, da webmaster e da archivisti e, come partner esterni, da un rappresentante della BNF²¹, da un rappresentante della Direzione degli Archivi di Francia (incaricata della missione presso la direttrice) e degli Archivi nazionali (CAC, programma Constance).

Gli archivisti del ministero hanno saputo approfittare dell'occasione. In effetti, mediante alcune correzioni, il sistema di gestione sviluppato all'interno dal servizio informatico corrisponde ai bisogni formulati dalla missione degli Archivi nazionali e dal CAC per la raccolta.

¹⁹ Institut national de l'audiovisuel.

²⁰ Institut National de Recherche en Informatique et Automatique.

²¹ Bibliothèque nationale de France.

In modo schematico, le operazioni devono svolgersi nel seguente modo:

- *preliminarmente la missione ha redatto un quadro di gestione gerarchizzato dei tipi di documenti;*
- *con un sistema di filtro, dal momento in cui un webmaster mette un documento in linea, i documenti destinati agli Archivi Nazionali sono trasferiti su un server dedicato all'archiviazione (in applicazione del quadro di gestione);*
- *l'archivista può convalidare la selezione operata automaticamente dal computer e scegliere di conservare o di eliminare i documenti immagazzinati sul server "archivi". I metadati possono essergli forniti dalla base di dati che il webmaster deve obbligatoriamente indicare per poter impostare i suoi documenti (la lista dei metadati ritenuti è quella proposta dall'ATICA);*
- *quest'operazione di archiviazione è attualmente in fase-test.*

La DGAFP-DIRE dei Servizi del Primo ministro poggia sull'esperienza in corso presso il Ministero del lavoro con delle specificità locali. Il lavoro cooperativo è in particolare uno strumento molto utilizzato da questa direzione.

La riflessione in corso porta a:

- *l'elaborazione di uno studio funzionale all'archiviazione;*
- *"l'esportazione" eventuale della base di dati sviluppata all'interno dal Ministero del lavoro verso altri ministeri per convenzione;*
- *lo studio dei costi (base dei dati, personale, materiale);*

In conclusione molte sfide restano da rilevare:

- *l'archiviazione delle pagine dinamiche;*
- *la concezione e l'elaborazione di strumenti di ricerca, le tecniche di conservazione.*

VIII. Conclusioni

Questo Manuale pratico dovrebbe permettere agli archivisti di dare prime risposte alle loro domande e a quelle dei loro servizi produttori.

Sarà seguito in breve tempo da altre opere di raccomandazioni e studi comparativi, sui supporti, sulla conservazione dei dati tecnici (piano, documenti catastali, archivi di architetti), dei documenti sonori e audiovisivi, sui piano di classificazione e sui quadri di gestione.

Quest'opera, oltre alla sua versione stampata, è in linea sul sito web della Direzione degli Archivi di Francia all'indirizzo: <http://www.archivesdefrance.culture.gouv.fr>. Questa versione sarà regolarmente aggiornata in funzione dell'evoluzione delle conoscenze in questo campo molto instabile dell'archivistica e in funzione delle osservazioni e delle nuove esperienze degli archivisti sul campo.

IX. Allegati

IX.1. Peso medio dei documenti elettronici e capacità dei supporti

Nome e misura approssimativa	Capacità di immagazzinamento dei supporti	esempio
Kilobyte	<ul style="list-style-type: none"> ▪ 10 linee di testo; ▪ ASCII = 1 kilobyte; ▪ 1 pagina Internet di testo in HTML con immagini statiche = 20 kilobyte; ▪ 1 pagina di un periodico elettronico in PDF = 50-500 kilobyte. 	1
Megabyte = 1.024.000 byte	<ul style="list-style-type: none"> ▪ 1 immagine digitale in JPEG = 50 kilobyte-1 megabyte; ▪ 50 articoli di 10 pagine in HTML = 1 megabyte; ▪ 2-20 articoli di 10 pagine in PDF = 1 megabyte; ▪ 1-20 immagini in JPEG = 1 megabyte; ▪ 1 immagine digitale in TIFF = 50-500 megabyte. 	1 dischetto = 1,47 megabyte; 1 CD-Rom = 650 megabyte.
Gigabyte = 1.024.000.000 byte	<ul style="list-style-type: none"> ▪ 5-50 pagine in TIFF = 1 gigabyte; ▪ 200-2.000 pagine in PDF = 1 gigabyte. 	1 cassetta Super DLT = 100-200 gigabyte
Terabyte = 1.024.000.000.000 byte	<ul style="list-style-type: none"> ▪ 200-2.000 giornali di 340 articoli di 10 pagine su un periodo di 10 anni in PDF = 1 terabyte; ▪ 50-500 libri di 100 pagine in TUIFF = 1 terabyte; ▪ 2.000-20.000 libri di 100 pagine in PDF = 1 terabyte. 	

IX.2. I mestieri legati all'informatica: esempio del Ministero dei lavori pubblici

A titolo di esempio affinché gli archivisti si rendano conto della diversità dei soggetti che intervengono e dei partner che potranno sollecitare o con i quali essi avranno da fare per gli aspetti tecnici, ecco ciò che avviene in un ministero, quello dei lavori pubblici.

Sei famiglie di raggruppamenti vi sono state identificate nel 1999 (risultanti sia dalla “maîtrise d’ouvrage”, o “mo”, sia dalla “maîtrise d’oeuvre”, o “me”).

Politica d’informattizzazione

- “Maître d’ouvrage” di applicazione
- “Maître d’ouvrage” di sistema di informazione
- Conduttore di operazione del dominio (funzionale): perizia e attenzione informatica, raccomandazioni ministeriali sull’evoluzione in corso (sintesi delle attenzioni trasversali applicata al dominio)
- Direttori e sotto-direttori d’amministrazione centrale (due sotto-direttori provenienti da due differenti direzioni).

Perizia

- Consiglieri di diritto e regolamentazione informatica
- Esperti di appoggio (rete strati bassi, strati alti, sistemi di gestione, materiale, cablaggio e telecomunicazioni, multimedia); effettuano una attenzione tecnologica; assistono tecnologicamente i supporti tecnici e i gruppi di sviluppo.

Studi e sviluppo

- RPMO (rappresentante permanente della “maîtrise d’ouvrage”).
- Capi di progetto di “me”.
- Amministratori di dati (“mo” o “me”) che garantiscono la coerenza dei dati veicolati sulle reti in termine di contenuto e di modelli evitando ambiguità, polisemia, ridondanza e che permettono la divisione fra differenti applicazioni (in particolare essi aggiornano e verificano la coerenza del dizionario dei dati utilizzato in un progetto).

Gestione e produzione

I mestieri della gestione necessitano sempre più delle conoscenze sui mestieri dell’amministrazione delle reti.

Si tratta all’occorrenza di:

- assicurarsi del buon funzionamento delle applicazioni sulla rete (sorveglianza del traffico, posto disco, tempi di risposta, risoluzione dei guasti, interventi contro dolo, gestione dei diritti d’accesso);
- assicurarsi delle procedure di soccorso delle applicazioni, di salvaguardia e di sicurezza dei dati (parametraggio dei diritti d’accesso, prevenzione contro il dolo, cambiamenti in caso di anomalia tecnica);
- assicurare l’informazione e assicurarne la coerenza;
- eventualmente gestire gli strumenti condivisi (stampanti, lettori condivisi);

- studiare l'evoluzione tanto in termini di materiali che di software per consigliare gli sviluppatori e i capi di progetto.

Funzione supporti (assistenza agli utilizzatori):

- diffusori di un'applicazione incaricata di promuoverla e di metterla in opera nei servizi (il che comprende l'installazione delle applicazioni degli utilizzatori);
- tecnici di supporto tecnico, con assistenza telefonica (cellula informatica);
- corrispondenti informatici.

Management dell'informatizzazione

- Webmaster: interazione con comunicazione e documentazione.
- Formalizzatori/ingegneri della conoscenza: suscettibili di avvicinarsi ai bisogni e alle pratiche degli utilizzatori modellizzando le conoscenze. Ciò presuppone conoscenze di tecniche di conversazione, di modellizzazione (di tipo sistemi esperti), di comunicazione uomo-macchina.

IX.3. Alcuni principi di base per valutare il costo dell'archiviazione elettronica

Non si tratta qui di fornire un metodo rigoroso per stabilir costi riguardanti l'archiviazione di documenti in formato elettronico, il che sarebbe un'utopia, ma di cercare di dare a un utilizzatore i mezzi per stimare il bilancio globale di un sistema di archiviazione elettronica:

- la cattura dei dati da archiviare;
- la realizzazione dell'archiviazione propriamente detta;
- la consultazione dell'archivio;
- la migrazione da una generazione tecnica di sistema di archiviazione a un'altra generazione.

IX.3.1. La cattura

Pur non volendo parlare propriamente dei costi dell'archiviazione, conviene richiamare qualche elemento riguardante i costi legati alla cattura.

Prima di tutto bisogna distinguere due tipi di cattura:

- la cattura di documento già in formato elettronico (per esempio la posta elettronica);

- la cattura di documento in un'altro formato (principalmente in formato cartaceo o di microforma).

Quando i dati sono già in forma digitale, esistono due costi principali:

- il costo di analisi dell'applicazione emittente dei documenti da archiviare;
- il costo della ricezione.

Il primo costo può essere molto variabile, perché l'applicazione emittente può essere già molto conosciuta (per esempio una gestione di compatibilità) o essere unica (gestione di una procedura amministrativa specifica di un ministero). È evidente che, per un'applicazione conosciuta, il tempo di studio è molto breve (una giornata), mentre, nel secondo caso, questo può arrivare fino a 2 o 4 mesi di studio.

Per quanto riguarda le spese legate alla ricezione, sono sia di telecomunicazione, sia di posta. Bisogna aggiungere costi legati ai controlli della ricezione. In funzione del tipo, possono andare da qualche decina di euro al mese a qualche migliaio.

Per i controlli, bisogna fare dei test in grandezza vera: apertura di un pacco, verifica del contenuto (numero di CD o di cassette, presenza di etichette, ecc.); conviene poi moltiplicare il tempo passato per il costo orario delle persone incaricate di fare questi controlli, costo che può essere elevato.

Per quanto riguarda la digitalizzazione, il quadro seguente dà una scala approssimativa dei costi. Questi costi sono legati principalmente:

- ai condizionamenti;
- ai formati;
- allo stato dei documenti;
- all'esistenza di colori;
- al bisogno di risoluzione;
- all'indicizzazione indispensabile alla gestione e alla ricerca dei documenti; i costi in questo campo varieranno enormemente se si trattano automaticamente i dati con l'aiuto di programmi di lettura automatica di documenti (LAD) o di riconoscimento ottico di carattere (OCR) o se li si tratta manualmente, il che può essere necessario se la qualità dei documenti è insufficiente per un trattamento automatico.

Nei due casi (documenti di origine elettronica o documenti ottenuti con digitalizzazione), il ruolo della compressione è fondamentale.

La scelta di un metodo di compressione e in seguito di un tasso di compressione è quasi sempre un compromesso tra il guadagno di spazio di archiviazione e la perdita di qualità dei documenti dopo la compressione.

Grosso modo, si può dire che classicamente i tassi di compressione dei metodi di tipo telescopica, permettono compressioni dell'ordine da 2 a 7 per documenti di qualità "fotografica", fino da 10 a 20 (per documenti di scarsa qualità).

Bisogna anche notare il caso particolare dei documenti elettronici corrispondenti a applicazioni aventi formati di edizione ripetitivi, per esempio i sistemi di fatturazione.

Esempio di un'applicazione che produce 50.000 fatture al giorno che contengono 30 kilobyte di messa in pagina e 1 kilobyte di dati (nome e indirizzo del cliente, quantità consumata durante il periodo, montanti HT, TVA, TTC). Se si immagazzinano le fatture documento per documento, si hanno, al giorno, 30 kilobyte per 50.000 cioè 1,5 gigabyte. Se invece si immagazzina da una parte il fondo della fattura (30 kilobyte) e dall'altra i dati utili (1 kilobyte per 50.000), si hanno, al giorno, 50 megabyte da immagazzinare. Il tasso di compressione è allora di 30, senza alcuna perdita. Ciononostante, bisogna descrivere in modo dettagliato come ricostituire il documento a partire dai dati e dal fondo di pagina.

IX.3.2. L'archiviazione

Il costo dell'archiviazione può essere scomposto in tre elementi principali:

- attrezzatura (materiale e software): il loro costo sarà molto differente se si utilizza, in funzione dei bisogni, un personal computer con un masterizzatore di CD-R o un grande sistema sotto UNIX con una libreria automatica di manipolazione di supporti di archiviazione;
- personale (personale di gestione e personale di sviluppo e/o di manutenzione secondo gli obiettivi);
- media o supporti di conservazione: un CD-R di buona qualità costa 150 volte meno di una cassetta di alta capacità.

Esempio

- *per 100 gigabyte al mese da immagazzinare;*
- *su CD-R il prezzo è di 2 euro l'uno;*
- *30 minuti per il trattamento di un CD;*
- *materiale ammortizzato in 3 anni;*
- *manutenzione uguale al 15% del prezzo del materiale;*
- *costo orario dell'operatore è, carichi compresi, di 15 euro.*

Si ha dunque:

- *ammortamento 2.000 euro/3.600 euro;*
- *manutenzione 2.000 euro per 0,15 euro 333;*
- *costo CD 2.000 per 2 euro 4.000;*
- *tempo passato 30 minuti per 2.000 1.000 ore;*
- *cioè in salario (1.000 per 15 euro) 15.000 euro;*
- *spese generali (affitto locale, riscaldamento, ...) euro 6.000.*

Il che fa un totale di 27.000 euro l'anno, cioè 21 euro a gigabyte.

IX.3.3. La consultazione/archiviazione

Il costo di questa risorsa dipende principalmente dal modo di consultazione:

- *accesso a CD-R con un magazziniere su domanda dell'utilizzatore;*
- *accesso in linea poco frequente (alcune consultazioni al giorno);*
- *accesso in linea frequente (molte migliaia di consultazioni al giorno);*
- *grande volume di dati (molte centinaia di megabyte per ogni consultazione).*

IX.3.4. La migrazione

Il costo di una migrazione si compone principalmente di tre elementi:

- *studio di fattibilità;*
- *costo degli operatori;*
- *costo dei nuovi supporti.*

È molto difficile immaginare questi costi, poiché sono dei costi futuri a tre, cinque o più anni. Si può dare il costo della migrazione dei dati e dei supporti realizzati nel 1996 per i 6.000 file conservati presso il Centro degli archivi contemporanei di Fontainebleau: 3 milioni di franchi per la migrazione di 4.500 file.

X. Glossario delle sigle

[A proposito di OAIS, Reference Model for an Open Archival Information System, modello di riferimento concepito dal Comitato consultivo per i sistemi di dati spaziali (CCSDS) nel 1999, portato all'ISO nel 2001 (ISO/DIS 14721.2-CCSDS/650.0-R-2), si forniscono le seguenti informazioni:]

Il sistema di archiviazione OAIS assicura la marcatura specifica di uno spazio digitale. In effetti, presentando l'oggetto di archiviazione come pacchetto di informazione, l'OAIS determina due tipi di informazione:

- *l'informazione contenuta;*
- *l'informazione complementare alla perpetuazione o PDI.*

L'informazione contenuta e la PDI sono incapsulate e identificabili tramite l'Informazione di Impacchettatura. Il pacchetto che ne risulta può essere ritrovato grazie all'informazione di descrizione. Nell'informazione complementare alla perpetuazione, si trovano quattro categorie di informazioni relative a:

- *la provenienza;*
- *il contesto;*
- *il riferimento o username;*
- *la fissità che assicura la protezione da ogni modifica.*